

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGCTGAAGACTAACATTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCTCCTCAGCCCTGTAATTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTTATGTCGTTAAAGCAAGTTCATGCTGTAGTCAGGAAACGTTATCATCAA
TTAAACAAAGGCTGCCCTGTACTTGGAAACTGAGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGTTACCTTGGTATGGCTCAT
TATATATGTTGTCAGACATCCTTCCACCAAATGCAGCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGCAAGTGGCAATTGGGACTGATTTAGAACAGAAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGCTATGTCATTTCTTCTT
GGTTTCTGACTTACATTGCTGATTTCAAGAAAATTCTTACGGGTGGAAGCCAATTAC
TGGATTAACCCCTATGACACTGCACCTGCCATTAAACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTATTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWRLLLVIWCGVSALSMLTCSVLHSGNFGTDLHQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACCGTGGGGAGAGCCGAGTCCCGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCCTGGCCCGAGCGTGTGGCCTGGGAGTGGAGTGGAGGCAGGAGCCTC
CTTACACCTGCCATGAGTTTCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTG
GATTGGGTGGCTTTCTCATGCCAATTGTTAAAGACTATGAGATACTGAGTGTACAG
GTGATCTCTCCGTGACGTTGCACTTCTGCACCATGTTGAGCTCATCATCTTGAATCTTAGG
AGTATTGAATAGCAGCTCCGTTTTCACTGGAAATGAACCTGTGTGAATTCTGCTGATCTGG
TTTCATGGTGCCTTTTACATTGGCTATTATTGTGAGCAATATCCGACTACTGCATAAACAGA
CTGCTTTCTGTCTCTTATGGCTGACCTTATGTTCTGAAACTAGGAGATCCCTTCC
CATTCTCAGCCCCAAACATGGGATCTTATCCATAGAACAGCTCATGCCGGTTGGTGTGATTGGAG
TGAACCTCATGGCTCTCTTCTGGATTGGTGTCAACTGCCATACACTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCAGTGTGCAAACCATGGATATGAT
CATAAAGAAAAAGAAAAGGATGGCAATGGCACGGAAACAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAAGTGTACCTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAGCAGGCAGCTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTCAAGGGAAATATTAAATTCTGGTT
ACTTTTCTCTATTACTGTGTTGGAAATTTCATGGCTACCATCAATATTGTTGATCGAGTT
GGGAAACGGATCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCCAACACATTCTTCTCATTCTGTTGAAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGATCCGAATGAGTATGCTT
AGAATACCGCACATAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACTTGACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTAA
GATATAAGAGGGGGAAAAATGGAACCAGGGCCTGACATTATAAAACAAACAAATGCTATGGTAGC
ATTTTCACCTTCATAGCATACTCCTCCCCGTAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATTCCGTGTGGATATGAGGCTGG
TGTAGAGGCCAGAGGAGCCAAGAAACTAAAGGTGAAAATACACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTTAGGATTCCGTTAAGGTTCACATGGAAAAGGTTAGCTT
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCCATGGCCAACCTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQLFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRGVIGVTLALLSGFGAVNCPTYMSYFLRNVTDTDILALERLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSPFWGMIKSVTTSASGSENLTLIQQEVDALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSŠNVIVLLAQMIMGMY
FVSSVLLIRMSMPLERYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATACCAGTGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGGGCTCTGGTCAAGGCTTTCGCGTGAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCCCTTGTGACAAGGGGAACCAAGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCTGTAGGCTGGGACTAAGTTGGCCGGAAGGACCAAGTGAAACAGCC
TTGAAAGCTAGCTTGTAAACTTGAGCTATGGCTGGGAGATGGATTGCGTGTCACTCTAG
GATTAGCCAAACCCCAAGTGTGGAAAATGGGTGGGTGTCTGATTGGAAAGGTTCCAGTGA
GCCGACAGTTGACGCTATTGTTACAACACTCATCTGATACTTGACTAACTCGTGATTCCAGAA
ATTATCACCACCAAAGATCCCATATTCAACACTCAAACGTCAACACAAACAGAAATTATGT
CACTGACAGTACCTACTCGGTGCATCCCCCTACTCTACAATACCTGCCCTACTACTACTCC
CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAAATTGATTGTGTCACAGAAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTGGGAGCTGAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CAGCTGGTCTGGATTGCTATGTCAGGAGGATGTGAAGGCCTCCCTTTACAAACAAAGAAT
CAGCAGAAGGAAATGATCGAACCAAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACCTACCGTGC
GATGCTGGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCTTACCTGCCAGCTGGGAAATCAAAGGGCCAAAGAACCAAAGAACAGAAAGTCCA
CCCTGGTTCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCAACAAAGAGAAT
GCCCTCTCTTATTGTAACCTGTCGGATCCTATCCTCTACCTCCAAAGCTCCACGCC
TTCTAGCTGGTATGCTCTAAATAATATCCCACGGAGAAAGGAGTTTGCAAGTGCAAGGAC
CTAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTCTGGTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGTGGACCAAGGCTCTCTACTGATTCCGCAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTCTGAGGCCGGTAAGAGCAAAGAAT
GGCAGAAAAGTTAGCCCCCTGAAAGCCATGGAGATTCTCATAAACTTGAGACCTAATCTGTAAA
GCTAAAATAAAGAAATAGAACACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGGAATCTGTTAGAACACACACA
CTTACTTTCTGGTCTTACCAACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAATCTTATAAATTCTATTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGTGCAAGGTATTACACTCTGTAATTGAAATTATTCTCTAAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTCTCTGAGTTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTCTCTGAGACTAATCTTATTCTCTAATATGGCAACCATTATAACCTTAATT
TATTATAACATACCTAAGAAGTACATTGTTACCTATATACCAAAAGCACATTAAAAGGCC
ATTAACAAATGTATCACTAGCCCTCCTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATT
TGTGACAAAAATTAAAGCATTAGAAAATT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNNPKCGKNGVGVLIWKPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAASFNEAAGFGGVPALLVLALLFFGAAAGLGFCYVK
RYVKAFFTNKNQQKEMIETKVVKEEKANDSNPNEESKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCGCTCCCGCACCGCGGCCGCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCCAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGCCACCCCTGCTGTGCCCTGCTGCCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCACCTCGGCTCAGTCAGGCCGGCG
GCTCTCAGCTACCCGAGGAGGAGGCCACCCCTCAATGAGATGTTCCCGAGGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAAGTGAACCTGGCAAATTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCATGTGCACCGAGAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTACATCTGTTGGAGACGAAGAAGGCAGAAGGAGC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTA
CTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGAGGACCA
GCTGTGTGTCGGGGTCACTGCACCAAAATGCCACCAAGGGGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGCTGTGCTGCTTCAAGAGAGGCCGTGCTGTTCCGTGTC
ACACCCCTGCCGTGGAGGGCGAGCTTGCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGGCTGATGGAGCCTGGACCGATGCCCTGTGCAAGGCTGACCTCTGCGCAGC
CCCACAGCCACAGCCTGGTGTATGTCAGGCCGACCTCTGCTGGGAGCCGTGACCAAGATGG
GAGATCTGCTGCCAGAGAGGGCTGGAGGAGCTGGAGAGGACTGACTGAAGAGATGGCCTGGGGAGCCTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCTGACTGAAGAGATGGCCTGGGGAGCCGTGACCAAGATGG
CCGCGCTGCACTGCTGGGAGGGAGAGATTTAGATCTGGACCAAGGCTGTTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCAGGTGTTGCTTACATGGCTTCAAGGCTTCTTCCCTA
CATCTCTCCCAGTAAGTTCCCTGCTGCTTACAGCATGAGGTGTTGCAATTGTTAGCT
CCCCCAGGCTGTTCCAGGCTTCACAGTCTGGCTTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCAGATTATGGCTGCTTGCCTCACCAGTTGGCAGACAGCCG
TTGTTCTACATGGCTTGTGATAATTGTTGAGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTGGGAATGGAGAGGTGCCCTGTTGCAACATCAACCTGGCAAAATG
CAACAAATGAATTTCACGCCAGTTCTTCATGGCATAGGTAAGCTGTGCTCAGCTGTTGC
AGATGAATGTTCTGTCACCTGCAATTACATGTGTTATTGTCAGCAGTGTGCTCAGCTCC
TACCTCTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGGG
AGGGGGCATTGTTCTCGCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTGACTCTAACGCTAGTGTCT
CTCCACTACCCACACCAGCCTGGTGCACCAAAAGTGTCTCCAAAAGGAAGGGAGATGGGAT
TTTCTGAGGATGACATCTGGAATTAGGTCAAACATAATTCTACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGCCGTCTTAATGAAGACAATGAT
ATTGACACTGTCCTTGGCAGTTGCAATTAGTAACCTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAAGAACAGTACTTAGGTATTGTTAGGGCGAGGATTATAATGAAATTG
AAAATCACTTACCAACTGAAGACAATTATCAACCAACTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGAAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTCATGGACTGTTGCCACCATGTATTGTCAGGTTCTTAAAGTTAAAGTTGCA
CATGATTGATAAGCATGCTTCTTTGAGTTAAATTATGATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAATCACTCAACTGCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECI IDEDCGPMYCFASFQYTCQPCRGQRMLCTRSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCACFQRGLLFPVCTPLPVEGELCHDPASRLLDITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAGTCCAATAAACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCGTGCCTGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGAAATCTATGGGGTCTGGGCTCTGGACCCCTT
AACTGGGTACTGGCCCTGGGCCATGCGTCCTCGCTGGAGCCTTGCGCTCCTCTACTGGGCTT
CCACAAGCCCCAGGACATCCCTACCTCCCCCTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCGGGTATCTT
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTGTCTCAGCAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAAGGGTGGTGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTGGGAAGCTGCTGGT
GGTCGGAGGGCGTGGGGCTCTGCTCTCTCTCCGGTCGATCCGGGCTGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCTAT
GTCATGCCAGCGGCTCTCAGCGTTTCGGCATGTGTGGACACGCTCTCCTCTGCTTCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGGCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCTTACAGGT
CTCCATTGTGGTAAAAAAAGTTTAGGCCAGGCGCGTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCCGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGCCAACATGGT
AAACCTCCGTCTCTATTAAAAACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCA
GCTACTCGGGAGGTGAGGCAGGGAGAATCGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAA
AAGATTTATTAAAGATATTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLFFFGLLGVLSFFFFSGRI PGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGGCTGCCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCCCTGCCCTCTGCCGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTCCTCTCCCTGGGGTGTGGTGTCCA
TCATTATGCTGAGCCCAGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGTCCCTGCAGGCCACATCGACTGTGGCTCCCTGCTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCCTCTCTCTTTCACCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTGGTCTTTAAGTCCGTATCCCTGGTGGGCC
ACCGTGGGTGCCCTTACATCCCTGACGGCTCCCTCACCAACATCTGGTTCTACTTCGGCGTGTGG
CTCCTTCCCTTCATCCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCGAGCG
GGCTGGGCAAGGCCAGGGAGTGCAGATTCCCGTGCCTGGTACGCAGGCCCTTCTTCACTCC
TTCTACTTGCTGTCGATCGCGCCGTGGCGCTGATGTTCATGTTACTACACTGAGCCAGGGCTGCC
CGAGGGCAAGGTCTTCATCAGCTCACCTCACCTCTGTGTCTGCGTGTCCATCGTGTCTGC
CCAAGGTCCAGGACGCCAGCCAACTCGGGTCTGCTGCAGGCCCTCGTACACCCAT
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAAACAGAAATGCAACCCCCATTGCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCCATCATCTTCCCTGTGACCCCTTTCATCAGTCTGCGCTCCTCAGACCACGGCAGGTG
AACAGCCTGATGCAAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGG
GGCAGCCTGTGAGGCCGGCCTTGACAACGAGCAGGCCGTCACCTACAGCTACTCC
ACTTCTGCCCTGGTGTGGCCTCACTGCACGTCTGATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACGCCGTGTGGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCCAACCCTGGTGCCTCTCGGCTCGGTGACAG
CAGGCCATCTGGTGCCTCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCC
CCCCACCAATCAGGCCAGGCTGAGCCCCACCCCTGCCAAGCTCCAGGACCTGCCCTGAGGCC
CTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCTGCAGAGCCCCATCCCCCGCC
ACCCACACGGTGGAGGCTGCCCTCCCTCCCTCCCTGGTGTGGCCTACTCAGCATCTGGATGAA
AGGGCTCCCTGTCCCTCAGGCTCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACTCCCAC
TGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGACCTGCCCTTCC
GACTTCGTGCCCTACTGAGTCTAAGACTTTCTAATAACAAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHAWNQRWLGKAE
ECDSRAWYAGLFFFITLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNNSGLLQASVITLYTMFTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIIFLLCTLFISLRSSDHRQVNLSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMIWTATVWVKICASWAGLLYLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCGTTAAGGTGCTTCTCTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTGC~~GA~~ATATCCATTCCATCAACCCCACACAA
CTCATGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGACCTTGA~~CT~~TATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTGATATATTCTCTGGCAGTTT~~G~~TAAAGTGTAAACTTGATATGCTGTGCAG
ACTGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCAAGGGCTTTGGCTATGTGCTGCCATCATT~~C~~ATT
CTTGCC~~T~~GGATTGAGACGTGGTCC~~T~~GGATTCAAAGTGTACCTCAAGAACAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTTTATTCCCCTCTGAATCGAACAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCAC~~TT~~TTAGAACATGAGTACTACTTTGTTAAATGTGAAAAACCC~~T~~CACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGTTTATTGAACAGCTAAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTG~~GG~~CTGGTAAGGTAATGT~~C~~ATGATT~~C~~ATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTAGTAAGCAAGATACTTTTATTCAATT~~C~~AC
AGAATGGAATTTTTGTTCATGTCTCAGATTATTTGTTACAGTTTAAAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTATTCTGTTGCATTATGTG~~T~~ATGGC~~T~~GAAGTGTGG
CTTGCAAAAGGGAGAAAGGAATTGCGAACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTTCTGATTGTTCTGAAATGTTCTAAATACTCTATTGAAATGC
ACAAAATGACTAAACCATT~~C~~ATATCATGTTCTGCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLQPQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCCTGCCCCCGCCCTCCCTGCGCCGCCCTCCCGGGACAGAAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCGCTCTGCTACTGGCCCTGGGCTGGGTGCAGGGCTGCCAT
CCGGCTGCCAGTGAGCCACAGACAGTCTCTGCACTGCCGCCAGGGGACCCACGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGCTGTACGTCTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTGCGGGCTGCGGGCTGCACTGCTGGACCTGTACAGAACAGATGCCAGCC
TGCCCAGCGGGCTTCCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGCCAACAGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCTGCGGCCCTGAGCGCCTCACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTCGACACGCTGACCGCCTCCTGGAGCTCAAGCTGC
AGGACACAGACTGCCAGCTGCCCGCTGCGCTGCCGCCCTGCTGCTGGACCTCAGC
CACAAACAGCCTGGCCCTGGAGGCCGATCCTGGACACTGCCAACAGTGGAGGCCTGCC
GGCTGGCTGGGCTGCAGCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCTCAGGGCTGACG
TGGATGTGTCCGACACCCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCTCAGGGCTGACG
CGCCTGCCGCTGGCCGAACACCCGCAATTGCCAGCTGCCAGCTGCCAGGACCTGCCGCC
TGCCCTGCCAGGAGCTGATGAGCAACCTAACGCTGCCAGGCCCTGCCCTGGGAGACCTCTGGG
TCTTCCCCCGCCTGCCGTCTGGCAGCTGCCCAACCCCTTCAACTGCCGTGCCCCCTGAGC
TGGTTGGCCCTGGTGCAGAGGCCACGTACACTGCCAGCCCTGAGGAGACGCCGCTGCC
CTTCCCGCCAAGAACGCTGCCGGCTGCTCTGGAGCTGACTACGCCGACTTGGCTGCCAG
CCACCACCAACAGCCACAGTGCCACACAGGCCCTGGTGGGAGGCCACAGCCTTGCT
TCTAGCTGGCTCTACCTGGCTAGCCCCACAGGCCGCAACTGCCAGCTGCCACCTGCC
CACTGCCCAACCGACTGTAGGGCTGTCCCCAGGCCAGACTGCCACGCCACCTGCC
ATGGGGCACATGCCACCTGGGACACGGCACCTGGCTGCTGTGCCCCGAAGGCTTCACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGACACGGCCAGCCCTACACCAGTCAGCC
GCCACCACGGCTCCCTGACCTGGCATCGAGCCGTGAGGCCACCTCCCTGCCGTGGGCTGC
AGCGCTACCTCCAGGGAGCTCGTGCAGCTCAGGAGCTCCGTCTCACCTATCGCAACCTATCG
GGCCTGATAAGCCGCTGGTGAAGCTGCCACTGCCCTGGTGCAGCTGCCCTGGTGCAGTACACGG
GCTGCCGCCCCAACCCCAACTTACTCCGCTGTGCTATGCCCTGGGCCCCGGGTGCCGAGG
GCGAGGAGGCCCTGCCGGGAGGCCATACACCCCCAGCCGCTTCAACTCCAAACCACGCC
CAGGCCGCGAGGGCAACCTGCCCTCATTGCCGCCCTGCCGCGGTGCTCTGCC
GCTGGCTGCCGTGGGGCAGCTACTGTGCGGGGGCAGCTGCCAGCTGCCCTGGTGCAG
ACAAAGGGCAGGTGGGGCAGGGCTGGGACTGGGAGCTGGGAGGTGAAGGTTCCCTGGAG
CCAGGGCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGGGCTGTGAGTGTGAGGTGCC
ACTCATGGCTTCCCAGGGCTGGCCTCCAGCTCAGTCACCCCTCCACGCAAAGCC
ACTATCTAACGCAAGGCTACATCTAACGCAAGGCTACATCTAACGCAAGGCTACATCTAAC
GAGAGAGACAGGGCAGCTGGGGCAGGGCTCTCAGCCAGTGAAGATGCCAGCCCCCTCC
ACACCCAGTAAGTCTCACTGCCAACCTGGGATGTGCTGAGACAGGGCTGTGACCAAGCT
GGGCCCTGTTCCCTGCCACCTCGGTCTCCCTCATCTGTGAGATGCTGTGGCCAGCTGACGCC
CTAACGCTCCCAAGAACCGAGTCCTATGAGGACAGTGTCCGCCCTGCCCAACGTGAGTC
CCTGGGCAAGGCCGCCCCCTGCCATGTGCGGTACGCATGCCCTGGTCTGCC
TCCAGGGCGGCCCTGGGGCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGG
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAGGAACAAAAGAAAAGTGGAAAGG
TTTAGGAACATGTTTGTCTTTTAAATAATAATTATATAAGAGATCCTTCCATTATCT
GGGAAGATGTTTCAAACTCAGAGACAAGGACTTGGTTTTGTAAGACAAACGATGATGAA
GGCTTTGTAAGAAAAAATAAAAGATGAACTGTGAA

FIGURE 16

MCSRVPLLLPLLLLLALGPGVQGCPSCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGQLQDDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLLELKLQDNELRALPPLRPRLLLLDLSHNSLLALEPGILDGTANVE
ALRLAGLGLQQQLDEGLFSRLRNLDLVDSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSLQALPGDLSGLFPRLLAARNPFCVCPPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGLIEPVSPSLRVGLQRYLQGSSVQLRSLRLYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPGLPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAQDKGQVGPAGPLEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGCGAGGCCGGCGTGGCTGAGTCGTGGCAGAGCGAAGGCAGCTCATGCG
GGTCGGATAGGGTGACGCTGCTGCTGTGCGGTGCTGAGCTGGCCTCGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTCAAAGAATCTGA
ATTAGAATCCTTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTGACCGCATTGAAGGCACAGCACATGGGAGCCCTGCCACTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGGAAATGAAAATCCTTAATGAAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCAGTGAACCATA
CCAAAGCCCTGGAGAGAGTGTCAATGCTTTATTGGTGAATTGCCCCACAGAAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTTCTGTATGCCCTGGACTGGTGTAAATTCAAGTCAGGCAAAGGCTTTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATAGCCACATGGTTGGTAAAGTAGACTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGTTAGCGGCCACAACCTTTCAAGCTTCAATGGATAT
CAGATTGCTTGTATTAAGACCAAATATTCAAGTGAACCTCCTTCAAATTCTGTAAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTGGTGGAGTCCACAATTCTTAAATGATTAG
TTGGCTGATTGCCCTAAAAGAGAGATCTGATAAATGGCTTTAAATTCTGTAGTTG
GAATTGTCAGAATCATTTTACATTAGATTATCATAATTAAATTCTTAGTTCA
AAATTGTAATGGTGGCTATAGAAAAACACATGAAATATTACAAATTGCAACAAATGC
CCTAAGAATTGTTAAATTCAAGTGGAGTTGGCTGAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCATGATTGGCTGCTTCCATTATTCTGGTCAATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTCCATTGGCTAATTGTTACTTTCTTGCTAATTGG
AAGATTAACTCATTAAATAAAATTATGCTAAGATTTTTTTTTTTTTTTTTTTTTTTTT
AA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLSKTTLTSDESVKDHTTAGRVVAGQIFLDSEESSEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPNENKDYEPPKKVRKPALTAEIGTAHGEPCCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAAKRRQMQEAEEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLQPQNIQAAREMFEKLTEEGSPKGQTAALGFLYASGLGVN
SSQAKALVYYTFFGALGGNLIAMVLSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATTTAAGCCCATTCTGCAGT GGAATT CATGAACTAGCAAGAGGACACCATCTCTT
GTATTATA CACAAGAAAGGAGTGTACCTATCACACACAGGGGAAAATGCTCTTGGGTGCTAGG
CCTCTAATCCTCTGTGGTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGGAAACTTGGCAGCCAGAACTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCGCGTGTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTTCCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAAATGTCTCCAGTGTGGAGGTGCCTTGCAATCGTGGAGGGGCTATACTCCATCCAAA
TATGCAGT GGAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACACAATATGGAGAAGGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAGTCTCTCCCTAACACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATA CCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCAACAAATGTCTCCAGGCTATGA
AATTGGCCGATTCAAGAACACATCTCCTTCAACCCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGTCTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGTATCCCAGGGT
CCCTGCTCAAGTTCTTGAAAGGAGGGCTGGATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTGCCTGCTGGTGTGATGTAAGGAAATTGAAAGACTTGCCATTCAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGCCCCAGCATTACAGTAAC TTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEFYIEKSLDKGNKSYVNMD
LSPVVECMDHALTSFLPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTAACACGGACTCGGACACCGAAGGTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATACAATTGA
CATTCAAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCAGTAAATG
AGCAAGCAGTGAAGAAAATATTCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCAAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAACTTGCAGGAGCATT
TTCAAACCAAGACCTGTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAAACCTCAAAAGGACTTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGATGTCGAACAACGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCAAATTTTGAAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
AAAAAGGAGAGGAGCACAGATTCAAGCAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGA
ACATTTCCTTCAGGCATTACGGACCTTTTCAAATTCTGAATTCTTCATTCATGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACCTACAACCACATCTGATGTAGT
AGACAATCTGACCTTAATGGTAAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAGCCTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATATTCACGGTCTC
CTACATTTGATCTTAAACCTTACAAGGAGATTTTTATTGCTGATGGTAAAGCCAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCACTAAGTTCAATTGTTTACTATGTTCA
TGTTGAGCTAATACACAGATAACTCTTAGTGCATTACTTCACAAAGTACTTTCAAACATCA
GATGCTTTTATTCCAAACCTTTTCACTTCACTAAGTTGAGGGAAAGGCTACACAG
ACACATCTTCTAGAATTGAAAAGTGAGACCAGGACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGAGTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCAGTCTATTAAAAAATGAAAAGCAAGAATAGCCTTATTTCAAAATATGAAA
GAAATTATATGAAAATTATCTGAGTCATTAAATTCTCCTTAAGTGTAACTTTTAAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVRNLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVKSSCNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCGAGTCCAGCCGGACGAGCGACCAGCGCAGGGCAGCCAA
GCAGCGCCAGCGAACGCCGCCGCCGCCACACCCCTCGCGGCCCCCGCGCGCCCTGCCACCCCTCCCTCCCC
GCGTCCCCGCCCTCGCCGGCCAGTCAGCTTGCGGGGTCGCTGCCCGCGAAACCCCGAGGTACCCAGGCCGCCCTCT
GCTTCCCTGGGCCGCCGCCCTCCACGCCCTCCCTCTCCCTGGGCCCTGCCACCGGGGACCGTGGCTGA
CGCGAGGCCAGCTACTTTCGCCCCGCTCTCTGCCCTGCTGCCCTTCCACCAACTCCAACCTCTCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGGCCCTGGCCCTGCCGTAGGCCGCTTCCGTCCGGTCCAAA
GGTGGGAACGCGTCCGCCGCCGCCACCATGGCACGCTTGGCCCTGCCGTCTCTGCCACCTGGCAGTGCTC
AGCGCCCGCTGCTGGCTGCCAGCTCAAGTCGAAAGTTGCTCGGAAGTGCAGCGACGTCTTACGGTCCAAAGGCTTC
AACAAAGAACGATGCCCTCCACAGAGATCAACGGTGTACATTGAAAGATCTGCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGATGATTCAAAGACTTCAAAAGATCTGAGTCAATTGCAATCATTG
CAAGCTGCTTGTCTCACGTTACAAGAAGTTGATGAATTCTCAAAGAAACTACTTGAAGATGAGAATCCCTG
AATGATATGTTGTGAAGACATATGCCATTATACATGCAAATTCTGAGCTATTAAAGATCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGGAGAATGCTAAAGTGTGAGTACGGAGCAGCTGAAG
TTCCGCCCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATGTTGAGCAGTACGGAGCAGCTGAAG
CCCTTCCGAGATGTCCTGCCAAATTGAAGCTCAGGTACTCGTGTGAGCAGTAAACTACTGCTCAAACATCATGAGAGGC
TTAGCGGTTGCGGGAGATGCGTGAGCAAGGTCTCCGTGGTAAACCCACGCCAGTGTACCCATGCCCTGGTGAAG
ATGATCTACTGCTCCACTGCCGGGCTCGTGACTGTGAAGCCATGTTACAACATGCTCAAACATCATGAGAGGC
TGTGTTGCCAACCAAGGGATCTGATTGAACTGAAACATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTCAACATTGAATCGGTATGGATCCCATCGATGTAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAACTGCTCAGAAGGTTTCCAGGGATGTTGAGCCAAAGCCCTCCAGCTGGACGAATTCTCGT
TCCATCTGAAAGTGCCTCAGTGCCTCGCTCAGACCACATCACCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTTACTGATGTCAGGAGAACTGAAACAGGCCAGAAATTCTGGCTCCCTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGAGAAACGCCAATGAGGATGACTGTTGAAATGGGAAAGGCAAAGCAGGTAC
CTGTTGCAACTGACAGGAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTCGTCAAATCATGCCCTTCAGTGTGACCAAGATGAAAGATGCTACACATGGAAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAGAGTGCCTGAGAAAGGCAATGAGAAAGCCGACAGTGTGTTGCTCCGTCTGGGCA
CAGGCCTACCTCCTACTGTCCTGCTCTGGTATGCAAGAGAGAGTGGAGATAATTCTCAAACACTGAG
AAAAAGTGTCAAAAGTAAAGGACCAAGTATCACTTTCTACCATCTAGTGAATTGCTTTTAAATGAA
TGGACAAACATGACAGTTTACTATGTTGCACTGGTTAAGAAGTGTGACTTGTGTTCTCATTCAAGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGAGTTGTTGCTGCCCTGCTCCCCAACATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
TTTTTTCCAACTGTGATCTCGCCCTGTTCTACAAGCAAACCCAGGGTCCCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAATAGCTGTACAGAAGCAGGTTTATTTATCATGTTATCTTAAAGAAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKS KSCSEVRRLYVSKGF NKN DAPLHEINGDHLKICPQGST
CCSQEMEEK YSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKS LNDMFVKT YGH
LYMQNSELFKD LFVELKRYYVGVNVNLEEMLNDFWARLLERMFR LVNSQYHFTDEYLECVSKY TE
QLKPF GDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHC RGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMILMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQGCGPPKPLPAGRISRS ISES AFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKF WSSLPSNV CND ERMAAGNGNE DDCWNGKGKS RYLFAVTGNGLANQGNNPEVQV DTS
KPDILILRQIMALRVMTSKMKNAYNGNDV DFFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCAGGAAGGGGCCAAGAATGTGAGTGCAAAGATTGGTCCTGAG
AGCCCCGAGAAGAAAATTATGACAGTGCTGGGCTGCCAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAACAAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTGCCTTGTAGGAGCTCTG
AGCGCCCCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAAGACACTC
TTCTCTCCACCTCACTCTCCACTGTACCCACCCCTAAATCATCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTAACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTCCT
AGCTAGTGTCATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPKCDHFKGKGNVKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTCTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCGTGCAAAAGCTACTTCCCCTACCTGA
TGGCGTGTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTTCAGCCAG
ATAAAGGGGTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGCCAACCGGAGC
CAACTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGTTGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCATGGATGTGGTCTGCACACTGGTCTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTCTGGGAGCATGTGGCAGAACATGGAAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCACCTGAAACACATTGGGATGGCTGCCTCACAGAGAGACCTGGAAGGATCTTGA
GAACGCCAGTTCTCGAAATCCAAATGGAACGACAGCCCCCTCCCTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGAAAGGTGTCAAACAAATTTCCAAGCTCCAAGGCACACTATTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTCCACTGAGAGGGACTA
GCAGAAATGAGAGAACATTATGTACCACTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCTCGACAGTAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCCTC
CCAATGTTGTCCTTCCCTCGTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTAGGAACTGGTCACAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTGAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATTTTTAATAAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLTLPLHIMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSSKALICSFPSL
QLEQATHQPIYLPRLGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGTCCCCAGAAAGTCTCTCTG
CCACTGACGCCCCATCAGGGATTGGGCCTCTTCCCCCTCCTTCTGTGTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTTA
AGAAAGCTGGGAGATAGGAACAGAACAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTTATTAAA
GTGGTTGTTATGATTCTTACTAATTACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTCCCTGTGTTCAATGTTGTAAGATTGTTCTGTGTAATATGTCCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTC SQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCAGAGCTTCAACTATACCCACAGTCCAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCTTATTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTGATAATGATGTTGACACCCCTCACCAGAATTCTAAGTGAATCATGTCGG
GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTGCTTGGCATGATGTTACC
TTCAGATTCACTACCACCCCTCTGGTCACATTTCATTCATTGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAGGCATACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCCTCTGGGCTCTGGGGCTGTGCTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCTTCGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTGTTATTCACACAGAAGTAAAATGATCCTCCTGATCATCCCATCCTTCGTCTCTC
TCCATTCTCTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCAATTAAATCTGTGGTGAG
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTG
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAAATGCATATACTACAACGTCTATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTTGCCAAGAACGCAACTTACATCTATTAAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTTGCTGTTGATC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
GGGAAACAGAACTCCAGGCCATTGTGAGAAGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTCTAAGAGCCATTACAGAATAGAAGATGAGACCAGTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAACCCATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGLLFVCGLWWLYYDYNLDSIE
LDTERENMKCVLGFIAVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISAPFLFQPLWTFA
ILIFFWVLWAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMVMQNALKEQQHG
ALSRYLFRCYCFCWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFCGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTCTTGGACCTCTCCCTGTTCTCCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAGCCTAAGGGAAAAGAATATTCAATTCTGTGTGGT
GAAAATTGGAAAAAAATTGCCTCTCAAAACAAAGGGTGTCAATTCTGATATTTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTCCTGTTGCTGGTACTGGAGTAC
ATTCAAAACAAAGAACGCAAGAAGATTAAAGGCCAAGTTCAGTGTGCTCAGATCAACTG
GATGTCAAAGCCGGAAGATCATCGATCCTGAGTCACTGTGAAATGTCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGCCACTGACGTGATGATCCTACTCCAGTGTGTGGCCTGCCG
TACACAGTGGTGTGTTGATAATTCAAGGAGGAATATTGTTGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATGTTATCCCTACACGATGGAGAGAAATCCTT
TATCGCTTAGAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCAGCTTACATACATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACGTGACAGCCGGTCACTCTGATGCGAGTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTTGCACAGGCCATCCCTCTGCTGTTTACCAACAGCATCCCCAGACCAACATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTCAAGGCAAGATCCTCAGGAGCTGCCCTCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGGACTGTTCAAAAGAAGAATTGAGCACACAGTCTTGGAGGCCAGTATCCC
TGGGAGATCCAAACTGCAAATTGACTTGTGCTTTTAATTGATGGAGCACCAGCATTGGCAA
CGCGGATCCGAATCCAGAAGCAGCTCTGCTGATGTTGCCAAGCTTGTACATTGGCCCTGC
CGGTCCACTGATGGGTTGTCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGACTTCTAAT
GTAGGTGGGCCATCTCTTGTGACCAAGAACCTCTTTCAAAGCCAATGGAAACAGAACGG
GGCTCCCAATGTGGTGGTGGTGAATGGTGGATGGCTGGCCACGGACAAGTGGAGGAGGCTCAA
GACTTGGAGAGAGTCAGGAATCAACATTCTCATCACCATTGAAGGTGCTGTTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCAACTTGTCAAACAAGGCCGTGTCAGAACAAACGGCTTCAACTC
GCTCCACGTGAGAGCTGGTTGGCCCTCCACAAGACCCCTGCAAGCCTCTGGTGAAGCGGGTCTGC
ACACTGACCCGCTGGCCTGCAAGAACGACTCTGCACTGGCTGACATTGGCTTCGTCATCGAC
GGCTCCAGCAGTGTGGGAGGGCAACTTCCGACCCGTCCTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGCGCATCGGGGGCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTGGGTTCCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGGACCCAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGGCCAACAGAGGAAGTTAATGATCTCATCAGCAGGGAGGCTTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCCATAGGGCTTGGCTGGGCT
GCCCAAGAGGAGCTAGAAGTCATGCCACTCCCCGCCAGAGACCACTCCCTTGTGGAGCAGA
GTTTGACAAACCTCCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC
AGCCTCGAACTTGAATTCAAGAGCAGGAGCAGACCAAGTGTGTTACTAATGACGTGTT
GGACCACCCACCGCTTAATGGGACGACGGTGCATCAAGTCTGGGAGGCCATGGAGAAC
AAATGTCTTGTATTATTCTTGCATCATGCTTTTCAATTCCAAAATGAGGTTACAAAGA
TGATCACAAACGTATAGAATGAGCCAAAGGCTACATCATGTTGAGGGTGTGGAGATTTCAT
TTTGACAAATTGTTCAAAATAATGTCGAATACAGTGCAGCCCTTACGACAGGCTTACG
AGCTTTGTGAGATTGTTAAGTGTATTCTGATTTGAACCTGTGAACTCTGTAACCCCTCAGCAAGTT
TTTGTCATGACAATGAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINC DVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGLDNGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTVPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPADPGIQQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPILMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPN VVVVMDGWPDKVEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKT LQPLVKRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFR TVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINF ALEQL
FKKS KPNKRKLMILITDGRSYDDVRIPAMA AHLKG VITYAIGVAWAQEELEVIATHPARDHSFF
VDEF DNLH QYVPRI IQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGTGGCTGCCTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGCTTTTCTGTTTCAGAGTAGTTCACAAACAGATCTGAGTGTAAATTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTTCATCTGGAATTCCACAGTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTGCTCTACGTGGTGCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTCTTGAGCTGCTGTGGATGCCCTCGGCTCTGGACTGTCCCTGGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCTGCTGTCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTGAACGGATGTAATTCTATGAGTATGAGCCGATT
CAGACAAGACTTCACCTCACACTCGAGAGCATTCAAACACTGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCCAACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGTCT
TGGTGGGGATATGAGGTTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAATGTTGGC
ATTGTCCTTAGAGGATGAAACACCTCTTATGGTACATAATCCGACAAGATTAGACACATATAATA
ACCTGACCTTGAAACCATTATGCCATTCAAGGTGGTAACTGAGTTTGCCTCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCATCAAACTGGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTTCACAGGTATCCTCTAATTGATAATTATTCTATAGAGGATTACCAAAAAACCCATATT
CTTACCAAGGAGTATCCTTCAGGTGTTCCCTCCACTGCACTGGGTTGGGTTATATAATGTCCAGAGAT
TTGGTGCCAAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAGTGAACATTCAATTCCAGAACAGACACAAATTTCTTCTATAGAATCC
ATTGGATGTCGCAACTGAGACGTGTATTGCAAGGCCATTGGCTTTCTTCAAGGAGATCATCACTTT
TGGCAGGTATGCTAAGGAACACCACTGCCATTATTAACTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGAAAGTGTAAATAAGTAGGTACTGTGAAAATTCACTGGGAGGTCACTGTGCTGGCTT
ACACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG
CCCTCAAAAGATGATATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAAAGG
ACCAAAACAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAAGGGTACTGAGTTAAAGCTCA
CTAGGCTGAAAAACAAACAAATGTAGAGTTATTGAAACAATGTAGTCATTGAAAGGTTTGCTGA
TATCTTATGTGGATTACCAATTAAAAATATATGTAGTTCTGTGTCAAAAACTCTTCACTGAAGTATA
CTGAACAAAATTTCACCTGTTTGGTCATTATAAAACTTCAGTCAAGATGTTGCACTGAGTTATT
ATTATTAAAAATTACTTCACATTGTGTTTAAATGTTTGACGATTCAACAAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTTCAGTTACTTAACGTGATCAGTTATTGATAACATCACTCCA
TTAATGTAAAGTCATAGGTCAATTGCAATATCAGTAATCTCTGGACTTTGTTAAATTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIAMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNNSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCGCCGCGAAGGATGGAGCTGGGTGCTGGACGCAGTGGGGCTCACTTTCTTCAGCTCCTTCATC
TCGTCTTGCAAGAGAGTACACAGTCATTAATGAAGCTGCCCTGGAGCAGAGTGGAAATATCATGTCGGAGTGCCTG
AATATGATCAGATTGAGTGCCTGCCCCGGAAAGAGGGAAAGTCTGGGGTTACCATCCCTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCATGATCCACCCAGGTTGATCCTTGAAGAACTGCAAGAGCTGCCAAATGGCTCATGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTACTGTGCAAGAGTGCAGGCCAGGGCTGGTACGGAGGAGACTGCATGCA
GCCAGGGTCTGCCAGGCCAAAGGGTCAGATTGTTGGAAAGCTATCCCTAAATGCTACTGTGAATGGGACATTGATGC
TAAACCTGGGTTGTCATCCAACTAAGATTGTCATGTTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCGTGAATGGAGACAACCGCGATGGCAGATCATCAAGCGTGTCTGGCAACGAGCGGCCAGCTCCATCCAGAGCATAG
GATCCCTCACTCCACGTCCTTCCACCCAGGGTCAGATTGTTGACGGTTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCTCATCCCCCTGTTCCATGACGGCACGTCCTGCTGACAAGGCTGGATCTTACAAGTGTGCCTGGCAGGC
TATACTGGGAGCGCTGTGAAAATCTCTTGAAGAAAGAAACTGCTCAGACCCCTGGGGCCAGTCATGGGTACAGAAAA
TAACAGGGGCCCTGGGTTATCAAGCAGGCCATGCTAAATTGGCACCGTGGTCTTCTTTGTAACAACTCCATGTT
TCTTAGTGGCAATGAGAAAAGAAACTGCCAGCAGAATGGAGAGTGGTCAAGGAAACAGCCATCTGCATAAAAGCTGCCGA
GAACCAAAGATTTCAGACCTGGTGAGAAGGAGAGTCTTCCGATGCAGGTTCAAGGGAGACACATTACACCAAGCTAT
ACTCAGGGCCTTCAGCAAGCAGAAACTGCAGAGTGCCTTACCAAGAAGCCAGCCCTTCCCTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTCATCACCCTCTACCGCCGGCTGGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCATCTGCATCCCTACTGCGGAAAATTGAGAACATCACTGCTCAAAGA
CCCAAGGGTTCGCTGGCGTGGCAGGCCATCAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTCTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGCTGCCACTGTGTTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTGGGAAATTCTACCGGGATGATGACCGGGATGAGAACACATCC
AGAGCCTACAGATTCTGCTATCATTCTGCATCCAACTATGACCCATCTGCTGATGCTGACATGCCATCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCGAGTCCAGCCATCTGCCCTGCTGCCAGTGGGATCTCAGCACTTCCAGGAG
TCCCACATCACTGTGCTGGCTGGAATGTCCTGGCAGACGTGAGGAGCCCTGGCTCAAGAACGACACACTGCGCTCGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGAGGAGCAGCATGAGGACCATGGCATTCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGTGGGAAACCACTGCCCTCTGATATCTGCACTGCAGAGACAGGGAGCATGCCCTGTCAGGAGCTGGGG
CGAGCATCTCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGTGGAGCTATGATAAAACATGCAGGCCACAGGCTCTCA
CTGCCCTACCAAGGTGCTGCCCTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACCTCTGAGAAG
TGTTTCTGTATATCCGCTGCTGAGTGTGTCATGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGCTGTGAACTTGGCT
GTGCCAGGGCTCTGACTTCAGGGACAAAATCAGTGAAGGGTGAGTAGACCTCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAGATGCCAGGGCTTGCAGAGTAAGTTCTTCAAAGAACGACATACAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCAACTTCTGAGTATACGAATGCCATCAGCTTGCAGGCCAGGGAGATCTGGCTCATGAG
GCCCTTTGAGGCTCTCAAGTCTAGAGAGCTGCCCTGGGACAGGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCC
TGTGTACATGGCCACAGTACAGTCTGGCTCTTCTCCCATCTTGTACACATTAAATAAGGGTTGGCTCT
GAACTACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCYEYDQIECVCPEGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNQEWGKQPIKACREP KISDLVRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRLGSSRTCLRTGKWSGRAPSCIPICGKIEINITAP
KTQGLRWPWQAAIYRRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREKTIQSLQISAIILHPNYDPILLADIAILKLLDKARISTRVQPICLAASR
DLSTS FQESHITVAGWNVIADVRSPGFKNDSLRSGVSVVDSLLCEEQHEDHGIPVSVDNMFCA
SWEPTAPS DICTAETGGIAAVSFPGRASPEPRWLMGLVSWSYDKTCSHRLSTAFTKVLPKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCATACATCCTCTCATCTGAGAAATCAGAGAGCATAATCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCATAATCTTGTATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTTGGCTGG
TTGGGGCCCTTGTAGCTGACAGAAGGGCCAGGGAGAATGCAGCACACTGCTGGAGAATGAAGGCGCTTGTGTTGC
TGGCTTGCCCTGGCTCAGCTGCTAACTACATTGACAATGTGGCAACCTGCACACTCCTGTATTCAAGAACTCTGTA
AAGGTGCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCAGCTCACAGATGGCTGTCCAGACGGCTGTGAGCC
TCACAGGCCAGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCCTGGCTAGACA
ACCCCTGCCCTACGTGTCCCTGGCAGAGGACGGCAGGCCAGCAATCAGCCAGTGGACTCTGGCCAGCAACCGAACTA
GGGCACGCCCTTGAGAGATCCACTATTAGAACAGATCAATTAAAAAAATAATCAGAGCTTGTGTTGAACTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCAGCAGGGCAGGGAAATTCTGAAAACACCAACTGCCCTGAAG
TCTTCCAAAGGTGTACCCACTGATTCCAGATGGTGAATTACAGCATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCAGAACCCACTGGTCCATATCATTATCCAACACATTATCGTGTGGGG
TGATGCCAGAGACGGCCGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACACTACGCTGTGCGTCTCTGCCAGGCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAAGTCCGCA
GCAGGAACAATGGCAGGGCCCGGATGCCCTACAGACCCCGAGATGACAGCTTCTATGTGATTCTCAACAAAGTAGCC
CCGAGGGCAGCTTGGAAATAAAACTGGTGCAGCAAGGGTGGATGAGCCTGGGTTTCTCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGGAGGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCCAAAAGTGCGGCTCATCTGATTCCAGGCAAGTGAAGACGTCACCTCGTGTCCCGCAGGTTCGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCAGGGGAGAGGAGCAACA
CTCCCAAGCCCCCTCCATCTACAATTACTTGTCAATGAGAAGGTGGTAAATATCCAAAAGACCCGGTGAATCTCTCG
GCATGACCGTGCAGGGGGAGCATCACATAGAACAGGTTGCTATCTATGTCACTCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTGTTGAATGTGGATGGGTCGAACAGAGGTGAGCC
GGAGTGAGGCACTGGCAATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGGCCAGCAGCCCTGGACTCCAACCACAAACATGGCCCAACCCAGTGACTIONGGTCCCTC
GGGTCACTGTGGCTGGAAATTACACGGTGCTGTATAACTGTAAAGATATTGTTACGAAGAAACACAGCTGGAGTC
TGGGCTTCTGCATTGTAGGAGTTATGAAGAATACATGAAACAAACCTTTTCTCAATCCATTGTGAAGGAA
CACCAAGCATAATGATGGAAGAATTAGATGTGGTGAATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAAATTACTCTAACTATTGTTCTGGCTGGCACTT
TTTATAGAATCAATGATGGTCAGAGGAAACAGAAAATCACAAATAGGCTAAGAAGTGAACACTATATTATC
TTGTCAGTTTATATTAAAGAAAGAACATTGTAAGGAAAGTGTGAGGAAAGTGTGATCATCTAAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCCAAAAAAATTAAAACACTAGTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATATTCTATTCAATAAAAGCCCTAAACAACTAAAATGATTGATTGTATACCCACTGAAATT
CAAGCTGATTAAATTTAAATTTGGTATATGCTGAAAGTCTGCCAGGGTACATTATGCCATTTTAATTACAGCT
AAAATTTTAAATGCTGAGAACAGTTGCTTCAACAAAGAATAATTTTCAAGAAGTTAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCA
S LTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARP
FERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINR
VDPSESLSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI
FNVLDDGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVNNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPG
VI SRDGR
IKTGDILLNVDGVELTEVS
RSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKS
IVEGTPAYNDG
RIRC
CGDILLAVNGRSTSGMI
HACIARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTCAGTTGTCATCAAGTTGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCTGCACTGAAGCAGAGAGATAGATATTATTACAGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGCTGACAGTGGTTGGGG
CACCAGTAACTACTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGAAGGGAAAAACTCTGACTAATGAAGCATTCCACGAAGAAGGTAGAACTT
GACAACGTCCCTCTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAG
CTTACAGAGGGTCGCCATCCTGTTCCCCACCGGAACAGAGAGAAAACACTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCACTCCACCAGGCTGAAGG
AAAAAGTTAATCGAGCCAAACTCTTGAATGTGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTAAACCTTACAAGTGTGAGGAG
CATCCCAAGCATTGGTGGTGGCAGGAACAGCACTGGGTACAGGTACAGTGGATATTGG
GGGTGTACTGCCCTAAGCAGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGAT
GGGGAGGCAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCGGCCCTG
CCTGAAGTGGTAAATATACAATGGCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACAGACTGGAGAACAGATGGTTGAGTAGTTGTTCTTAAAT
TAGTATCTGTGGAACACAATCCTTATATATCAACATCACAGTGGATTCTGGTTGGTGCATTGACCC
TGGATCTTGGTGAATGGAAACTGATTCTTGTGCAATAATTGGCTAGAGACTCAA
ATAGTAGCACACATTAAGAACCTGTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTCT
TAGCAGAGCTCCTGGTGAATGGAGTAGAGTATAAAACAGTTGTAACAAAGACAGCTTCTAGTCATTGAT
CATGAGGGTAAATATTGTAATATGGATACTGAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTTTAAATTGAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCACTTACAGTACAATCATCTGTGAAGTGGTGGTGTAGGT
GAGAAGGCCTCACAAGAGGGAGAAAAGGCAGCAATCAGGACACAGTGAACCTGGGATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCAGTGTGATGCCACCAGAGAATACATTCTATTAGT
TTTAAAGAGTTTGTAAAATGTTGACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTA LSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGGGACTGAGCTCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTGCCAGCCTGACCAGTGGCTCTGTTCCCACAAACAG
ACGGGACAACCTGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCAGCTGGATGCCAT
GTTCCAGAGGCAGAGGAGGCAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAAACCTACCTGCCCTGCCCCGTCCCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLTLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRHSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTAAAGAGCTGGTCGGT
TCCGGTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTGTACTCTATTGTCTG
GACCTCAACACAACCCCTTGTCAACCATAACGCCAGAAGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGATATACAGCTCATCAGCAGCCCTCCACCC
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGTATGACCTCATCTGCGTTGCCAGGAACCTGTCAGCAGAAACT
TCTCAAGCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCTGCTCAGTCTTTGTACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTGTCGGG
AAACTCTAACATATGCCCTATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAAATCCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGGAAATACCGAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYIILWQLTGSAAAGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEPDANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCCTCAAAACCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGAATTCAATTCCAGAATTCTCAA
CTTGCAGTGGTTTCAATGACTCTTGACCTCCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGC
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGAAATAAAATGTAAGTATCAGTAGTTGAAAAA

FIGURE 48

MTCCEGWTSCNGFSLVLLLGVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGCTGCAGACCCCCCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTGAACTGTGA
CATGGAGAGAGTGACCTGGCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCGCCATTGCTGGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGAGCACAGTCCTGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCTAACACTGGCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCACTGGACAAATATTGGGGAGGCCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCAG
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCATGCTCTGGAAACACTGGCAGCAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGGCTCCTGGCAGGG
GTGCCTGGCACAGTGGCTGGAAACTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGACCAACACTCAGGGAGCTGTGGCCAGGCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGTGGCGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGCAGTGA
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCGGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAAATGAAGCCCCGGAGCGGGGAATCTGGGATTCAAGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTCTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCGAGCTGGGAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACTCTGAGACGGTCTCTGGGATGTTAACCTTGACACTTCTGGAA
GAATTAAATCCAAGCTGGTTCATCAACTGGATGCCATAAACAAAGGACAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACAGATTGGATGGAGCCCCCACACTCCCTCTTAA
AACACCACCCCTCTCATCACTAATCTCAGCCCTGGCCTTGAATAACCTTAGCTGCCCAACAAA
AA
AA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQQTREAVGTGVRQVPGFAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQQGNPGGLGTPWVHYPGNSAGSGMNPQGAPWGQGGNGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGSSNSGGSGSQGSSGSGSNGDNNNGSSGGS
SSGSSSGSSGGSSGGSSGSSGSGSNGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNHKPGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGACAAGCTGCTCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCTCAGACCGTGGCAATGTCCTGGCTACTCCCTGCTGCTGGTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTGGACCTATGCCCTCTATAACAACTGCCGCCGGCTCCAGTGTTTCC
CACAGCCCCAAAACGGAACTGGTTTGGGTCACTGGGCTGATCACTCCTACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCCAT
CATCCCCTTCATCGTTTATGCCACCCTGACACCATCCGGTATCACCAATGCCCTAGCTGCCA
TTGCACCCAAAGGATAATCTCTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCTTCCATTCAACATCCT
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCCTCATGACCTTGACAGTCTA
CAGAAATGCATCTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCTTGTAGAGAAAAGAGCCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGCGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGTCGCACCCCTCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGTCTGTACAAACCTTGCAGGCCACCCAGAATACCAGGAGCGTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGCC
AGCTGCCCTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATCACAAACCAACTGTGTTGGCCGATCCTGAGGTCTACGACC
CCTTCCGTTTGACCCAGAGAACAGCAAGGGAGGTCACCTCTGGCTTTATTCCCTTCTCCGCA
GGGCCCAGGAACTGCATCGGGCAGGCGTTGCCATGGCGGAGATGAAAGTGGCCTGGCGTTGAT
GCTGCTGCACTCCGGTCCTGCCAGACCAACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCAGATGACTTTCTGAC
CCATCCACCTGTTTTGCAATTGTCATGAATAAAACGGTGCTGTAAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVGSWLLARI LAW TYAFYNNCRRLOC FPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITN ASA AIA PKDNL FIRFLKP
WLGE GILLSGGDWKSRHRRMLTPAFHFNILKSYITIFNKSANIMLDKWOHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFD SHCQER PSEYIATILELSALVEKRSQHILQHMDFLYYL SHDGR FHRAC
RLVHDFTDAVIRERRRLPTQGIDDFFKDKAKSKTLDFIDVLLSKDEDGKALSDEDIRAEADTF
MF GGHD TTASGLSWVLYN LARHPEYQERC RQEVQELLKDRDPKEIEWDDLAQLPFLTM CVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDP EYDPFRFD PENS KGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTCTGTTGCACCTACCCGTGTTCTGCCT
TTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTGGCATTGACGTGG
TACAGCCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTGC
ATAATTCATGGCCAGTTTATGAAGCTTGAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGCTTACAGACATGTGCCTTATCTTGCAGTGTTGCTT
GTGATTGAAACATTGAGGGTACTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVIFPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMKDNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVDEFTEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRNGNGINVLKIAPESAIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATA
CGTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTTTGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTTGGACTATAATGCCAGCTCAGAGACCTTGGCGGTGAGGCTCCCGATGGTCCCC
CAGCCCACAGTGGCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACCTTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAACATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGTTG
CTCTCTTCTTGGCCATCAGCTGGCACTCTGCCTCTAGCCCTACCTGATGCTAAAT
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAAGATATGACCTAGTTTATTTCTGGGAGGAATGAATTCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAGAAGCAGAGCAGAAGGCTCAATATGAACAAGATAAAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTCTGTAACAGACAAGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCAAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGGAGGACAGGAGTAGTGCATGTTCTGTCTGAAATTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAGTCTATCCCAACATATCCACATCTTATATTCCAC
AAATTAGCTGTAGTATGTACCTAAGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAGTTGAGAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTATAAATAAAACTGAGCACCTCTTTAAACAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVLFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFQOPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIIRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAACATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCTCCGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTTATGGGGTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACTCCAGGAGGACCATATG
GCGGTGAGCTCCGGGGGCCCTATGGTCAGCCACCTCAAGTTCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGCTACGGCTCTCAGCCGTGGAATTCCAGCAGTGGAAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCATTAGCTACACAGAGCTGCAGCAAGCTCTGCCC
AAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGAGAAGGACACAGCTGTACAGCAACATCCGGCTCAGCTCGAGGACTCGTCA
CCATGACAGCTCTCGGATGCTTGACCCAAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGTTCTTAGAGTGAGAGAAAGTATGTTGACATCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AAATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTGTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCGTGATGCCAGTGGTAGTGTCTCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTCTCCAAAGTGGAAAT
CTGACCAAGCAGTGAGAGAGATCTGTCTATGGACCAGTGGCTGGATTCTGCCACACCCATAAAT
CCTTGTGTTAACTTCTAGCTGCCCTGGGCTGGCCCTGCTCAGACAAATCTGCCCTGGCAT
CTTGGCCAGGCTCTGCCCCCTGCAGCTGGACCCCTCACTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGCACCTCTCCCTGCCAATACTTTTTAATTGCAATTTC
ATTGGGCCAAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTGPGPYGGAAPGGPYGQPPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQLVTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGGCCTCTGGGCCTGCTCCTGGCTGTCTTCATC
TCCCAGGCCTTTGCCCGGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACCTCGGGACC
AACTTGCCCTAGCTCGGACAACCTCCTCACTGGCCCCCTAACTCTGAACATCCGAGCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGGAGTTCTGCAGTGAGGTTGGCTCCATCGTGGGGCTGCCATG
GATTCCCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCCAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACCTCTTACCTCTCAGTGTGCGGCCCTCGCTCCGGGAGTGGCCCTT
TGCCTGGGAGTCTCTCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGGAGCCGGGGAAAATCCTTCCAAACG
CCCTCCCTGGTCTCATCCACAGGGTCTGCCTGATCACCCCTGGGTACCTGAATCCCAGTG
TGTCCCTGGGAGGGTGGAGGCCCTGGGACTGGTGGGAACGAGGCCATGCCACACCCCTGAGGG
ATCTGGGTATCAATAATCAACCCCCAGGTACCAAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCAATCTATACCCAGGTATCAATAACCCATTCCCTCTGGAGTTCTCCGC
CCTCCTGGCTTTCTTGAACATCCCAGTGGCTCCCTAATCCTCAAGCCCTAGGTTGCAGTG
GGGCTAGGCACGATAGAGGGAAACCCAACATTGGGAGTTAGAGTCCTGCTCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPAGGSQVRWPPSWGLPAMDSWPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNLSGAGGKILSQRP
PWSLIIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMHPPEGIWGINNQPPGTWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCTTGTCTCCTCTTGCACCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGCTTTCTGACAAATTCCCTCATGAGTCCAGCTTCCGGAA
TTGCTTGAAAAGCTCGCTCCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCTTCTTGGCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCGTCTTCAGCAGGCCACCC
CTGAGTGGCAATAAATAAATTCTGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLGGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATGCGGTAGGAGGGCGAGCGCGAGAACGCC
TTCCTCGCGCTGCCAACCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGCGAGCCTGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGTCCCTCTCCCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSSDGTLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENKHADIVIVAPPTLPGRDEPYTKQFTECGEKEY
IHFTPDLGGKKQNEYGPPGKLFVHEWAHLRGVFDEYNEDQPFYRAKSKKIEATRCSAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLLQTVENGSWGMVHFSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFIALLGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDAKSMAVLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNNDGVSYRYFTAYTENGRYSLKVRAGGANTARLKLRRPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGKVRQYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVI
VNFIESTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACGTCCCCGGGAGGGGTGA
CAACAGGTGTCATCTTTGATCTGTTGGCTGCCCTCTATTCAAGGAAAGACGCCAAGGTAATTGACCCA
GAGGAGCAATGATGTAGCCACCTCTAACCTCCCTCTGAACCCCCAGTTAGCCAGGATTTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCTCGGCTAACCTGTTGGAGGAGAAACCTTGTGGGCTGGCTCTTAGCA
GTGCTCAAGAGTGAATTGCTGAGGGAGAAAGAAGGAAAGGTTCCCTCTGTGTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTAACCTGGAGATTCACTGCTAGTATTGCTTCGCTGCAAGATCATCCTTAA
ACTAGAGAACGCTGCTGTGTTGTTACTCCAAGAGGAGAGAACTGTTCTAGAAGGAAATGGATGCAACAGCTC
CGGGGCCCAAACGCATGCTTCTGTGGCTAGCCAGGGAAAGCCCTCCGTGGGGCCGGCTTGAGGGATGCC
ACCGGTTCTGGACGCGATGGCTGAATGATGATGGTCTGCCGGGCTGCTTGCCTGGATTCCCGGGTGGT
GTTTGCTGCTGCTCTGCTGTGCTATCTGCTCTGTACATGTGGCTCACCCCAAAGGGTGAAGGAGGAGCAG
CTGGCACTGCCAGGGCAACAGCCACGGGAAGGGGGTACCCAGGGCTCCTCAGGACTGGGAGGAGCAGCAC
CGCAACTACTGAGCACCTGAAGCGGAGATGCCACAGCTCAAGGAGGAGCTGAGGAGAGGAGTGAGCACTCAGG
AATGGGCACTACCAAGCCAGCGATGCTGCTGGCTGGACAGGAGCCCAAAGAGAAAACCCAGGCCACCTC
CTGGCCTCTGCACTCGCAGGGACAAGGCAGAGGTGAATGCTGGCTCAAGCTGGCCACAGATGAGCAGTG
CTTTGATAGCTTACTACAGAAGGTGACCTGAGGACTGCCAACCCGAGGAGAACCTGTG
AGGAAGGACAAGGGGGATGAGTTGGTGAAGGCAATTGAATCAGCCTGGAGACCTGAACAATCCTGCAGAGAACAGC
CCCAATCACCGTCTTACACGGCTCTGATTCTAGAAGGATCTACCGAACAGAAAGGGACAATTGAT
GAGCTCACCTCAAAGGGACACAAACCGAATTCAAACGGCTCATCTTATTGACCATTCAGCCCATCATGAA
GTGAAAATGAAAAGCTCAACATGGCAACACGTTATCAATGTTATGTCGCTCTAGCAAAAGGGTGGACAAGTT
CGGCAGTTATGCAAGGAAATTCAAGGGAGATGTGCAATTGAGGAGATGGGAGTCCATCTACTGTTGTTACTTGG
AAAGAAGGAAATAATGAAGTCAAAGGAATACTTGAAGGAAACTTCAAGGCTGCCAACACTTCAGGAACCTTACCTTCATC
CACGCTGAATGGAGAATTCTCGGGGAAAGGGACTTGTGATGTTGGAGGGCTCTGGAGGGAGCAACGTCTTCTC
TTTTCTGATGTTGACATCTACTTCACATCTGAATTCTCAATACGTCAGGCTGAATACACAGCAGGGAAAG
GTATTTTATCCAGTTCTTCAGTCAGTACAATCTGCATAATATACGGCCACCATGATGCACTCCCTCCCTGGAA
CAGCAGCTGGTCATAAGAAGGAAACTGGATTGGAGAGACTTGGGATGACGTTGCACTGACCTTATGCAAGTATCTC
TTCATCAATATAGGTGGGTTGATCTGGACATCAAAGGCTGGGAGGATGTGACCTTATGCAAGTATCTC
CACAGCAACTCATAGTGTACGGACGCTGTGAGGACTCTTCACCTCTGGCATGAGAACGCTGCACTGGCAGAG
CTGACCCCCAGCAGTACAAGATGTGCACTGCAAGGCACTGAACGAGGACATCCCACGGCACTGGCATGCTG
GTGTCAGGACGAGATAGAGGCTACCTTCGCAACAGAAACAGAAAGACAAGTAGCAAAAAACATGAACTCCAGA
GAAGGATTGTGGAGACACTTTCTTCTTCTGCAATTACTGAAAGTGGCTGCCAACAGAGAAAGACTTCCATAAA
GGACGACAAAAGAATTGGACTGATGGTCAAGAGATGAGAAAGCCTGGCAGGAGATGGCTTCTCTGTTGGCTTTAACACAGA
AATACAAAATCTCGTTTGCCTCAGGAAACTGACCTGCAACCTGTGAAGTGTCTGACAAAGGAGAATGCTGTG
AGATTATAACGCTAATGGTGGAGGTTTGATGGTTTACAATACACTGAGACCTGTTGTTGTCATTGA
AATATTGATTTAAGAGCAGTTGTAAGGAAATTCTAGCATGAAAGGCAAGCATATTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTAGTTCTAGGAATGCTAAAGGAAACTGAGGAGGAGGAGATGGCTTATGCAAGTACT
AGTGAAGTACATTAAGTAAATAAGGACGAGAAAAGAACGAAATGATGTCATATTCCCAGAGAT
TAACCAAAATAATCTGTTTGCCTTAAACTGTCCTTCTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT
TTTTCTGTTGAGTTAGTGTGTTTATTAATTACACTTGCAGGCTTACAAGAGGACAAGTGGCTAC
ATTTTTATATTGTTAAGAAGATACTTGAGATGCTATTGAGACCTTCAGTCAAGCATCAAATTGATGCCATAT
CCAAGGACATGCCAAATGCTGATCTGTCAGGACTGAATGTCAGGACATTGAGACATAGGGAGGAAATGGTTTGTACT
AATACAGACGTACAGATACTTCTCTGAAGAGTATTTCGAAGAGTATTGCAAGGAGGACAACACTGAGGAGAAAAGAAAATGAC
ACTTCTGCTTACAGAAGGAAACTCATGAGCTGGTATATGTCATGTCACCTTAAAGTCAGAACCCACATT
CTCTCAGAAGTAGGGACCGCTTCTACCTGTTAATAAACCAAAGTACCTGTCAGGAAACAAACATCTCTTCT
AAAACAGGGTGCCTCTGGCTTCTGGCTTCAAGAAGGAAATGGAGAAAATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGAAGGAAGTTTGCTACATGTTATCCACCCAGGGCAGGTTGGAG
TAACTGAATTATTGTTAAATAAGCAGTTACATAGAGTGGTTCTCATTCTGAAATTTGCAAGTATTGATTCATT
CAAACATTGTTAAATAAGCAGTTACATAGAGTGGTTCTCATTCTGAAATTTGCAAGCAGCAC
ATGCAAGTGTGGTGTGTTAAAGCATTGATTGATTGACTGGTAGTTATGAAATTAAATTAAACACAGG
CCATGAATGGAAGGTGGTATTGACAGCTAATAAAATGATTGATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDRDELVEAIESALETLNNA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMVKNEKLNMAN
TLINIVPLAKRVDKFRQFMQNREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFQY
NPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCOYRSDFINIGGFDDIKGWGGEDVHLYR
KYLHSNLIVVTPVRLFHLWHEKRCMDELTPEQYKMCMOSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCTGGGATCCA
GAAACCCATGATACCTACTGAAACACCGAATCCCCTGAAAGCCCACAGAGACAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCTGTCTAGTCCTAGTCCTCAAATTCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATGGCCAGCCTCTTACCCGTAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCCTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTTACCCGTATCTGGGTGGACTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACCAAGATCAACAG
TGAAGCCACATTGAGAGCTCCACATTGTACATTATGACTCTGATCCCTATGACAGCTTGAGT
AGGCTGCTGAGAGGCCCTAGGGCTGGCTGTCCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTTCAACCTAACAGAGAGCTGCTCCCCAACAGCTGGGAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAGAGTGTGCTCTGGACAGTTTTATAGAAGGTCCCAG
ATTCAATGAAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCCACAGAACAGGGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTCTGCCTTCTCTGGCTGTTATTCTATTGCTAGAAAGATTGGAAGAACAGGGCTGGAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCCCTCAGATAC
CATGGATGTGGATGACTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCTCCCTGGACATCTCTTAGAGAGGAAT
GGACCCAGGCTGTCACTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTGTTAGTTGCAGGGAAAGTTGGG
ATATACCCCAAAGTCCTCTACCCCCCTCACTTTATGGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTGATCAATATATTTGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMSLGVGILVGCLLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGACGGCAGTCCCTGTGTC
TCTGGGGTTTGCCTAACCTGAAACATCACCTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGTCTCAAGGAGTTAAAGTTACTTACACTGTGAGTATTACATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGTCCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGTACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTCAAGGCTAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTCTTTCTTT
TCTGTGATGGGCTATTCCATCTACCGATATCCACGTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTTATGAAATGAATTGACAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTGGATGATTCTAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGGTATGCTTCGCAATTGATGAAATTGGACTCTGAAG
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGAGGTGTCACACAAGGAACATTGGACTCGCAGGCAGCGTGGCAG
TCTGGGCCGCAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGTCACTGGGA
TCCCCAAACTGGCAGGCTGTATTCCCTCGCTGTCCAGCTCGACCAGGATTAGAGGGCTGCG
AGCCTCTGAGGGGATGGCTGGAGAGGAGGGCTTCTATCTAGACTCTAGAGGCCGCTATA
CCAGACAGGCCACCAGGAGAAATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTTCTTTGCCCTTGTGAAACAAAGTGAG
TCACCCCTTGATCCAGCATAAAGTACCTGGATGAAAGAAGTTTCCAGTTGTCACTGAGTGT
CTGTGAGAATTACTTATTCTTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGCTGGCGTTGTTATGCAGAGAA
AGCAGTCATAAATGTTGCCAGACTGGGTGCAAGATTATTCAAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSLNTKSNRTWSQCVTNHTLVLWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPPQEEEVKHTMLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOQEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEPSTTLVDWDPQTGRLCIPSLSFDQDS
EGCEPSEGDSLGEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMCN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCAC TTGCTGCCCTCTGACAC
CTGGGAAG**ATGGCCGGCCCGTGGACCTTCACCCCTCTCTGTGGTTGCTGGCAGCCACCTTGATC**
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCAAGTACACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAAGTGCAAGTGGC
CCCACCCGCGCTGGCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACTGCTGTA
TAAGCTCTCCTCCCTGGTGAACGCCCTAGCTAACAGGTCAACCTCTAGTGCCTACCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCTGATCGAGGGCTCCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCATTCACTGGGGCCAAGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCACCCTGGACAAACATCCCG
TTCAGCCTCATCGTAGTCAGGACGTGGTGAAGAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCAATGGTCCCTGTGGACTCTGTGCTTCTGAGAGTGCCCATGGCTGAAGTCAGCATGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTAAGATCCTAACTCAGGAC
ACTCCCCAGTTTTTATAGACCAAGGCCATGCCAAGGTGCCAACTGATCGTGTGGAAAGTGT
TCCCTCCAGTGAAGGCCCTCCGCCCTTGTTCAACCTGGCATCGAACGCCAGCTGGAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACCTCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTGATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGT**GAAGACTTGGATGGCAGCCATCAGGGAAAGGCTGG**
GTCCCCAGCTGGAGTATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANIQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPILDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTCTACTGAGAGGTCTGCCATGGCTCT
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTCTGGGCTTGGGACACTGGTGC
GCTGCTCCCAGCTGGAAAACAAGTTCTATGTCGGTGCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCAGTGTGACATCTATAGC
ACCCCTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGC
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGACAGTCTCTGCCAGGAATCCC
GAGCAAAGACAGAGTGGCGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGGATT
ATTCCCTGCTGCTGGAATCTCATGGGATCCTACGGGACTCTACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTCTTACTGGGCAATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATTCTGCTTCCCTGCTCATCCAGAGAAATCGCTCCAACTACTACGATGCCTAC
CAAGCCCAACCTCTGCCACAAGGAGCTCCAAGGCCGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGAAGAACCAAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAACAGTGGACAGCACCCGAGGCCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTGCTGCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTGGTTACCTGGGACTCCATCCCAAACCCACTAATCACA
TCCCACGTGACTGCCCTCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTTAGCTCATT
GCTGGGGATGGGAAGGGAGAAGCAGTGGCTTGCTGAGGTTGGCTTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAACTGAAATAAAACCATCCACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSAAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLLHGILRDFYSPLVPDSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACGCGTCCGCGCCTCTCCCTCTGCTGGACCTTCCTCGTCTCCATCTCTCCCTCTTC
CCCGCGTTCTCTTCCACCTTCTCTCTTCCACCTTAGACCTCCCTGCCCTCCCT
GCCACCGCTGCTTCCCTGGCCCTCTCGACCCCGCTCTAGCAGCAGACCTCCCTGGGCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCGTGTCTTCTGTCTCCCTCCGACTCCGCTCCGG
ACCAGCGGCCTGACCCCTGGGGAAAGGAGGAGGTGGTCCCGAGGTGAGGGTCTCTCCCTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTGAGCCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCGAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCTATGAGCTGTTCCCTCCGCTGCCAACCCAGTGTGTCCCTGCACTGC
ACAGAGGGCCAGATCTACTGCAGCCTCACAAACCTGCCCGAACCGAGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAACGCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGTGCTGGGAGAAAG
AGAGGCCGGGCACCCAGCCCCACTGGCCTCAGGCCCTCTGAGCTTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACGTCAAGATCGCCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCTTCCGTGCCTTCGGC
CCCTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTGCCCCAGAGG
ACAAAGCAGACCCCTGGCACAGTGAGATCAGTTCTACCAAGGTGTCCAAGGCACCGGGCCGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCCGTGCTTGCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACCTCCGACTGCTCGTGGCCCCACGAAGGTCACT
GGAACGTCTTCCATTAGCCCCAGACCCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTTATTATATTAATAAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGTPAPTGLSAPLSFIPRHFRPKAGSTTVKIVLKEKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARNPERGTALPTARWPPRRSLERLPSDPGAEGHGQSQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCCCTCTCTAATCCATCCGTACCCCTCCTGTC
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGAACAGCCTGTCAGGCCAGGGAG
GACGCAGCATTCTCTGTTCCCTGTCCTAAGACCAATGCAGGCCATGAAAGTGCCTTCTCAGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGCAGGGAGGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGCGATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGCCCTATGGTGCAGGATTAGTCCCAGTCTACTACCAGAACGCCATCTGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTCCCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCACAAGGACAGGATTGTC
ACAGACTCCAGGACAAACAGAGACATGCATGCCCTGGTATGTTGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCCGCATGCTCATCTGAGCCGAGGGTGGAAATCCAGGGTACAGATAG
GAGATACTTTTCGAGCCTATATCGTGGCACCTGGTACCAAAGTACTGGAATACTCTGCTGTGGCTA
TTTTTGGCATTGTTGACTGAAGATTTCTTCCAAATTCCAGTGGAAATCCAGGCCAGTGGACT
GAGAAGAAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCTGCCTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTGTCTCCGATCATGGTACTGGTCTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCGTTTATCAGCGTCTCCCCAGGACCCCACCTACAAAATAGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCCTCTCAACATAATGACCACTCCCTTATTTACCCCTGACATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAAGAGGCCCTTGGCAAAGGGCTCTGCAATCCCAGAGAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTTCCCTCCCCAGGGTCAAATGTAGGATGAATCACATCCCACAT
TCTTCTTAGGGATATTAAGGTCTCTCCAGATCCAAAGTCCCGCAGCCGGCAAGGTGGCTTCCA
GATGAAGGGGGACTGCCGTGTCACATGGAGTCAGGTGTCATGCCCTGAGCTGGAGGGAAAGG
CTGACATTACATTAGTTGCTCTCACTCCATGGCTAAGTGATCTTGAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTCATTATATTACACTTCAGTAAAAAA

FIGURE 84

MALMLSVLSSLKLGSQWQVFGPDKPQALVGEDAASFCLSPKTNAEAMEVRFTRGQFSSVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPPLISITGYVDRDIQLLCQSSGWFPRPTAKWKGPGQGDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRRAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEQELDWRRKHGQAEQELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGCCCTGGCACCTAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCCTCCTACCCCTCGCATGGCTGGATT
TACCTGGCCCAGTAGTCATGGCTACTGGTCCGGAAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGCAGTGTGGAGGAGACTGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGATGCCGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGAAATTATAAACATCACGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCAACCCGCTCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATTGGCACCAGCCTCACCTGTCAAGGTGACCTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTAACGTGCTCACCGCCTCAGAACATTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTATCTGTCACTCCCAGAGGGCCAGTCT
GCGCCTGGTCTGTGCAAGTTGATGCAAGCAATCCCCGCCAGGCTGAGCTGAC
GAGGCCTGACCCCTGTGCCCTCACAGCCCTAAACCCGGGGTGCTGGAGCTGCCCTGGTGAC
CTGAGGGATGCAGCTGAATTACACCTGCAGAGCTCAGAACCCCTCTGGCTCTCAGCAGGTCTAC
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGA
CTCAGGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTGCTCATCTCGTTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCTGGGAGATA
ACGGGCATAGAGGATGCAAACGCTGTCAGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTGGGAGAAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCGCTCTCAGTGGGAGAGCTCCAGTATGCATCCCTCAGCTTCAAGATGGTGA
AAGCCTTGGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTAAAAGCACACAGGCTTAGAGTCAAAGTATCTAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAGACAGACAAATTCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHVPFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNSSL
SLPEGQSLRLVCAVDADSNPPARLSLWRGLTLCPSPQSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSAQGPLTEPWAEDSPPDQPPPASARSSVGEGELOYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTTATCTACCAAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAGTGCACGGT
GGCGATCGCTGGTCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGC
CAACTACAACACCTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCTCCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGC
CGGTGATCCCTGTGGCTATGATTTGGCAGGCCAGAAAACAGCATCTTATTACTCACCCAT
GGCCAGCGGAAATTCACTGCAGGATTGTTCACTCAGGTATTTAATAACGAGAGAGCAGCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGAGGAG
GATACTTCCAGAGGCCAGTCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCGTGGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCAGACCTCTCCTCCAAACCATGAGATCCAAAGGATGGAGAA
CAACTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIATTRGWTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNPKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTNDGPVIPVYDFGDAQKTASYYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGC~~GGGG~~AGACTTCAGGAGTCGCTGTCTGAAC~~TTCC~~AGCCTCAGAGAC
CGCCGCCCTTGT~~CCCC~~GAGGCCCATGGGCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGT~~CAC~~CTTATTCTGGTCCC~~GG~~ACAGCAACATACAGGCC~~TG~~
CCTGCCTCTCACGTT~~CAC~~CCCCGAGGAGTATGACAAGCAGGACATT~~CAG~~CTGGTGGCC~~CG~~GC~~T~~CT
CTGTCACCC~~T~~GGGC~~C~~CTTTGCAGTGGAGCTGGCCGGTT~~C~~CTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTACTGTAGTGCATCCGTGCC~~C~~TG~~C~~TTCTTCAT
ATT~~C~~GAGC~~G~~T~~T~~GGGAGTGC~~ACT~~AC~~G~~TATTGGTACATTTG~~T~~CTT~~C~~TGCAGTGCC~~C~~TTCCAGCTG
TC~~A~~CTGAAATGGCTTATT~~C~~GT~~C~~ACCG~~T~~TTGGGCTGAAAAAGAAACCC~~T~~CTGATTAC~~TT~~CA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTCGTATT~~C~~CTGGAAGAAGGAAG
GCATAGGCTTCGGTTTCCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTAGTGC~~T~~TGTAATAAAATGTTTG~~T~~AGTAACA
TTAAGACTTATATACAGTTAGGGACAATTAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALEFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGACGAGAGCGGAGGAGGAAGAATGCAACTGAC
TCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCTCCGGCTCAGAGGACCCGTAGCGTGATGACCACAGGGCCAGCCCCGGCCGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAACGTCCGCCCATGGCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTGGGAGGCTTGGGCATTCTGGGAGCCCCCAACGCCGAACCACAGCCCC
CACCCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGCAGTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCCAGTAAAGCTGTAG

AGTTCCACCAGGAACAGCAGATCTCATCGAAGCCAAGGCCTCCAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGGACCTCGTTGCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGT
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
CATAGTGATAACCCCTACTACCCATCTGGGTGACCCGGGAGGCCACAGAGGCCAGGGC

TGGAAGGACAGGCCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAAGGGTTGGGCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGTGGTCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGCACAGAGAGATGCTGG
TCCCCGAGGCCTGTGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAC
CCTTGGTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTCACTCAGTGTGGACA
GCCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGTTGG
GCCCAGAGGAGCTCCAGCCCTGCCAGTGTGGCCAGTACTCCCCCTGTACCAACCCATTGCTGATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTTGACAGATTGACCACTGTCTCCAGC
CAGGCCACCCCTTCCAAATCCCTCTTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATGG
GCACACCCCATCTTAAGCTAACGACAGGACGATTGTGGCTCCACACTAACGGCACAGCCCATC
CGCGTGCTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCCTGTGGAGCATCCATGTCCCG
GAGAGGGGTCCCTAACAGTCAGCCTCACCTGTCAAGACGGGGTCTCCGGATCTGGATGGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGCCAGACGATGTGCTGGATCTGTC
TGTGTTGCTGTGCTGTGGTGGGGAGGGAGGGAAGTCTTGAAACCGCTGATTGCTGACTTT

TGTGTAAGAATCGTGTCTGGAGCAGGAAATAAGCTTGGGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPNSRPMA
NLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTCGGCTGCGCCTTCATTGCCTCGGGCTGCGCTGCC
CTTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTCTGTCCTGTTGGTTATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCTATGTTCTGGCTTGGCTTGGATCATGA
GTGGAGTATTTCCCTTGTGAATACCCATCTGACTCCTGGGCCAGGCACAGTGGGCATTATCAT
GGAGATTCTCCTCAATTCTCCCTTATTCACTGCTGGCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCTTATCGTC
TCCTGACCCACCTGCTGGTGTCAAGCCCAGACCTTCATAAGTTCTTATTATGGAATAACCTGGCG
TCAGCATTATAATCCTGGTGTCAAGCCCAGACCTTCATAAGTTCTTATTATGGAATAACCTGGCG
AAGCCTGAAACTCTGCCTGCTGCCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT**A**
AACCTCAGGGAACCAAGCACTTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACGTGCCT
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHGDSQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSLYYGINLASAFIILVLMGTWAAGGSCRSLKLCLLCQDKNFLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGCTTAGGATCAC
TCGGTCATTACCACAGCTAAACCTGTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTCACACAACTTGGAGCCCAGGGCACTATCCTAAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAATCTGCCACCCAGG
GAACCCCGCAGGCCCTCCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGGCCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTTCAACTAAGCTGCCCGAATTGGTGATAATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTTCAGAAAATATGTTATGTAGAGAAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAAATCTGTCTTGAAATATAACATTATGCTGCCGGATGATATGCATAT
TAAAACATATTGGAAAACTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLILPLTQM
LTLPDPLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCTGCCCAACCCAGCCCAGGCCAGAGCCCCCTGGAGAAGGGAGCTCT
CTTCTTGCTGGCAGCTGGACCAAGGGAGCCAGTCTGGCGCTGGAGGGCTGCTGACCAGT
GTCCCTGCCTGGCTGCTTGTCTCCGCCCCAAGGCTCTCCCCAAGGGCCAGCTGC
AGAGCTGTCTGTGGAAGTCCAGAAAATGGTGGAAATTCCCTTATACCTGACCAAGTGC
CGCTGCCCGTGAAGGCCAGATCGTGTGCAAGGGACTAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTCCTGCTGGTGAACAGGGCCTGGACCGAGAGGA
GCAGGCAGAGTACAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGCTTGTGGGTCAC
AGCTGTGCTGTGACGTGAAGGATGAGAATGACCAGGTGCCCCATTCTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGTACCGAGGCTGGATCCCCCTCCCTCTGAGGGCTCAGACCG
GGATGAGCCAGGCACAGCCAACTCGGATCTCGATTCCACATCTGAGCCAGGCTCCAGCCAGC
CTTCCCCAGACATGTCCAGCTGGAGCCTGGCTGGGGCTCTGGCCCTCAGCCCCAAGGGAGC
ACCAGCCTTGACCAACGCCCTGGAGAGGACCTACCAGCTGTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAGTCTCCATCATAGAGGACCTGGGT
CCCTAGAGCCTATCACCTGGCAGAGAATCTCAAAGTCTTACCCGACCCACATGGCCAGGTA
CACTGGAGTGGGGGTATGTGCACTATCACCTGGAGGCCATCCCCGGGACCTTGAAGTGA
TGCAGAGGAAACTCTACGTGACCAAGAGACTGGACAGAGAAGGCCAGGCTGAGTACCTGCTC
AGGTGCGGCTCAGAATTCCATGGCAGGACTATGCGCCCTCTGGAGCTGACGTGCTGGT
ATGGATGAGAATGACAACGTGCTTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGTCACTAGACTGTCAGCAGGGATGCAGATGCCCGGCTCCCCCA
ATTCCACGTTGTTACGCTCTGAGCCTGAGCCTGAGGATGGGTAGAGGGAGAGCCCTC
CAGGTGGACCCCCACTTCAGGAGTGTGACGCTGGGGTGTCCACTCCGAGCAGGCCAGAACAT
CTCGCTCTGGTGTGCTGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTTGAG
TCGAAGTCGCACTGACAGATCATGATCACGCCCCCTGAGTTCATCACTTCCCAGATTGGCC
ATAAGCTCCCTGAGGATGTTGGACCCGGACTCTGGTGGCCATGCTAACAGCATTGATGCTGA
CTCGAGGCCGCTTCCGCTCATGGATTGGCATTGGAGAGGGAGACACAGAAGGACTTTG
GCCTGGATTGGGAGGCCAGACTCTGGGATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAGCTGGGGCCAGGCC
AGGCCCTGGAGGCCACGCCACGGTGAAGTGTGCTAGTGGAGAGGATGATGCCACCCCCCAAGTGG
ACCAGGAGAGCTACGGGCCAGTGTCCCCATCAGTCCCCAGGGGCTTTCTGCTGACCATC
CAGCCCTCCGACCCCCATCAGCGAACCCCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTCTCCGGGAGGTGCAACCGCCCAGTCCCTGCAAGGGGCCAGCTGGG
ACACCTACACGGTGTGTTGGAGGCCAGGATACAGCCCTGACTCTGCCCCCTGCCCC
TACCTCTGCACACCCCCGCAAGACCATGGCTGATGTTGGAGGCCAGCAAGGACCCCCGATCT
GGCCAGTGGCAGGGCTTACAGCTCACCTTGGTCCAACCCACGGTCAACGGGATTGGC
GCCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCAATTGGTGGAGGCCAGT
GAACACATAATCCCCGTGGTGGTCAAGCCACAATGCCAGATGTTGGAGCTCTGGTGTGAGTGA
CGTGTGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGGCCATGAAGGGATGCCA
CGAAGCTGCGCAGTGGCATTCTGTAGGCACCCCTGGTAGCAATAGGAATCTCTCATCCTC
ATTTCACCCACTGGACCATGTCAGGAAGAAGGACCCGATCAACCCAGCAGACAGCGTGC
GAAGGGCAGACTGTCTGATGGCCAGGCACTGAGCTCTAGCTGGAGCTGGCCCTGGCTCC
TCCCCCTGGGAGAGGCCAGCACCCAAAGATCCAGCAGGGACAGGACAGAGTAGAAGCCCTCCA
TCTGCCCTGGGGTGGAGGCCAGCATCACCACATCAGGCACTGCTGAGAGCTGGGCC
TATGGACTGCCCCATGGGAGTGCTCAAATGTCAGGGTGTGTTGCCAATAATAAGCC
CTGGGCTGGGCCATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLOVTLEMQDGHLWGPQPVLVHVKDENQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPPLELHVL
VMDENDNVPICPPRDPPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNILLVLAMDLAGAEGGFSSCTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGATATVTLVERVMPPPQKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLICIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNTVQRDWRLQTLNGSHAYLTLALHWEP
REHIIPVVVSHNAQMWMQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTIVAI
LIFTHWTMSRKDKPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGTACATTGCTGGAGGAAGCTAAGGAACCCAGGCATCCAGCTGCCACGCCGT
AGTCCAAGATTCTTCCCAGGAACACAAACGCTAGGAGACCCAGCTCTGGAAGCACCAGCTTA
TCTCTCACCTCAAGTCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGGCCACAAAGAGACAGATGAAGATGCAGAAAGGAATG
TTCTCCTTATGTTGGTCTACTATTGCAATTGAGCTGCAACAAATCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCAGGCCACCAACTCTGGTCCAG
TGTGACTCCAGTGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGAACCTCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCACAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTCAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAG
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCAGCACAGTCAGCACACCCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCAGCACAGTCAGCACAGTCAGCACACCCCTCC
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACCTTCCATAGTCATCTACTGCACTGAGTGTGAGGCAAAGCCTGGGGTCCCCTGGT
GCCGTGGGAAATCTCCATCACCCCTGGTCTGGTTGTGGCCGGCGTGGGGCTTTGCTGGC
TCTTCTCTGTGTGAGAAACAGCCTGCTCCCTGAGAAACACCTTTAACACAGCTGTCTACCCCT
CATGGCCTCAACCATGGCCTTGGGCTCCAGGCCCTGGAGGGAAATCATGGAGCCCCACAGGCCAG
GTGGAGTCTTAACCTGGTCTGGAGGAGACCAACTATCGATAGCCATGGAGATGAGCAGGGAGGA
ACAGCAGCCCCCTGAGCAGCCCCGGAGCAAGTGCCTGAGGAAAGGAAAGAGACCTGGCA
CCCAAGACCTGGTTCCCTTCATTCTCACCTTCTGCTTACAGACACTGGAAAGAGAATCTGAA
AATCTGAGAAGGTATTCTCACCTTCTGCTTACAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAATAACATCTCATCTAACACACAGACAAAGAGAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTGAGATGAACCTAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMKGKGNVLLMFGLLLHEAATNSNETSTSANTGSSVISSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTTSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGLVPWEIFLITLVVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHGLGPGPGENHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCCGT TACGGGATGAATT AACGGCGGTCCGCACGGAGGTGTGACCCCTA
CGGAGCCCCAGCTTGCACGCACCCCACTCGCGT CGCGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAAC TATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAGCTTGGAAATCAT
GGTGTCA TGGAAAGGGATTACTTACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAAGT
GATTATAACTGGGATGCATTGTCCTGGAGAAAGAAGTGTCA TTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTGGATTGGTGGCCATGCGAGCTGCTGCCTATAT
CTTCATTCA TAGGAAATGGAAGGATGACAAGAGCATTGCAAGACATGATTGATTACTTTGTG
ATATTCA CGAACCACTCAACTCCTCATATCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCACAAACGGTGGAGAGAAGAGAGAGGCTGCGTCCCTATCAAG
GGGAGAAGAATT TACCGGACAGAGTGTCA TCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTCAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGGTATTATAATCACCATTGAAATCTTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACCTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCTAACCTTCTAACGCTCAGATTTGCATGACTATGCGAATATTTCTACT
GCCATCATTATTGTTAAAGATATTTGCACITTAATTGCTGGGAAAAATATTGCTACAAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGTACGGGTGAAATAACTT
GGGCCAGAATATTAAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPMVALLETMFGVKVIITGDAFVPGERSVIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGWAMQAAAYIFIHRKWKKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFPREIHFHVRYPIDTLPLTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLOERIFGG
LEIIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCCACGGCTCTGCGCTGAGACAGCTGGCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTCTAGTGTAGATCAACCCACAGGAATA
TCCATGGCTTTGTCATTTGGTTCTAGTTCTAGGACGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCAGGGCTTGGTGGGAGGACGCCGTGTTCTCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCATGGAAGTGCAGGGCTTCAGGAATCAGTCCATGCTGGTC
CACCTCTACAGAGATGGGGAGACTGGGAATCTAACGAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAAGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCG
ACATCGCCTGTATGGGTGTCAGTTCCAGATTACGATGAGGAGGCCACCTGGAGCTG
CGGGTGGCAGCACTGGGCTCACCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCAGGCTGGTCCCCCAGCCCACAGCCAAGTGGAAAGGTCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTTGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGTGACAGACTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTTGGCCTGGCTTCTA
TTTACTCGGGTTACTCTGTTGCCCCCTGTTGTTGATGGGATGATAATTGTTTCTTC
AAATCCAAGGGAAAATCCAGGGCGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGC
TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGGTATGTTGAGGAGTGTGTCGGGATGACGTAGACAGGGGAAGAAC
ATGTGACTTGTCTCCAAACAATGGGTATTGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTACACGAGTAGGGGTCTCC
GGACTATGAGGGTGGGACCATCTCCTCTCAATACAATGACCAGTCCCTATTATACCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCTATATCCAGCATGCACTGACGAGAAAAG
GGGACTCCCATATTCAATGTCAGTGTCTGGGATCAGACAGAGAACAGCCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACGCCAAGGGAGAGTGCCTCCGACAGGTGGCCAGCTTCCCT
CCGGAGCCTGCGCACAGAGACTCACGCCCAACTCTCCTTAGGGAGCTGAGGTTCTGCCCC
TGAGGCCCTGAGCAGGGCAGTCAGCTTCCAGATGAGGGGGATGGCTGACCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGAGCGGAATAGACTCACATTAGGTTAGTTGTGAAA
CTCCATCCAGCTAAGCGATCTTGAAGTAAAGAGACAGAACGAACTGTAATCATGCTTCAGGTT
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAGAACGAACTGTAATCATGCTTCAGGTT
TGAGGGCACAGTGTGCTAATGATGTTTATATTATACATTCTCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTATACAAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAAGAGGAGGTAGGTTTCACTGATTCTATAAGGCCAGCAT
TACCTGATACCAAAACCAGGCAAGAAAACAGAAGAAGAGGAAGGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATAATT
AAGATGATATATAACTACTCAGTGTGGTTGTCCACAAATGCAGAGTTGTTAATATTTAAAT
ATCAACCAGTGTAAATTCAAGCACATTAATAAAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFQPTAKWKGPGQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQQPSPWRLASILLGLLCGALCGVVMGMIIIVFFK
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWVGVCRDDVDRGKNNVTLSPNNGYWVRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGTGGAACCTGGTTATGCCCTCGTCATCTCATATCCCTGATTGTCCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAAGAACCTACAATTACTATAGCACATTGTCAATTACAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTTAACAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAATCTCCATTAGGAAAGAATTGTCAAGTCTCAGGTTATCAAGTT
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCACTACTGAGGATCTGA
AACTGTAGATAAAATTGTTCAACTTGTACATGAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCAATTGCTGCC
ACACGAAGAAGTAAAACCTCTAGGTCAAGACTCAGGATCGTTGGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGATGGAGTCATCGCTGTGGAGCAACCTTAATTATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTACAACATATAAGAACCTGCCCAGATGGACTGCTTCCTGGAGTAACA
ATAAAACCTCGAAAATGAAACGGGTCTCCGAGAATAATTGTCCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGCAGAGCTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGTGACAGGATTGGACACTGAAAAT
GATGGTTACAGTCAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTACTCTGGAGGACCAGGTTAGTCAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTGTATAGAGTTACGGCCTGCGGGACTG
GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCTCATGGAACAGATAACATTTTTGTGTTTTG
GGTGTGGAGGCCATTAGAGATACAGAAATTGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAACTGTTGCTTGATGCATGTTCTTCCAGCTGTTCCGCACGTAAGCATTGCTTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCAATTGTCAGTCTAGAAGTTGTCAGAATTGACTTGTGACATAATTGTAAT
GCATATATAACAATTGAAGCACTCCTTCTTCAGTTCTCAGCTCCTCTCATTTCAAGCAAATATCCATT
TCAAGGTGCAAAACAGGAGTGAAGAAAATATAAGAAGAAAAAAATCCCTACATTATTGGCACAGAA
AAGTATTAGGTGTTCTTAGGAAATTAGAATATTGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAAATACCAATCATTCACTTCAATTAGGAAAGTATGGAACTAAGTTAGGAAAGTCCAGAAAGCAAG
ATATATCCTTATTTCATTTCCAAACAACTACTATGATAATTGTGAAGAAGATTCTGTTTTGTGACCT
ATAATAATTATAACAAACTTCATGCAATGTAATTGTTCAAGCAAATTAAAGCAAATTATTATTAAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAACACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKGRL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
IVSWGDECAPKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAAGAGCCTCCAGCTGAAGCCATGCAGCCCCCGCTCCCGCAAGAAGTCCCTG
CCCCGATGAGCCCCCGCGTCCGCTCCCGACTATCCCCAGCGGGCGTGGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTGCCCTGGGAGTAGGATGTGGTAAAGGATGGGCTCTCCCTT
ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCCTGCTACGCCCTCAA
TCTGCTTTGGTAATGTCATCAGTGTGTCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTCCT
GTGGTTCATCCGGTCACTGATTGCTGTTGCTGTTCTATCATTGCTGGGATGTTAGGATATTG
TGGAACGGTGAAAGAAATCTGTTGCTCTGCTGGACATATGAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGAATTTCAGAGAGAGTTAACGTGCTGGAGTAGTATATTCACTGACTGTTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTTAGAGAATTCCAGATGTCAAA
CAGGCCACCAGGAAGATCTCAGTGCACCTTATCAAGAGGGTGTGGGAAGAAATGATTCCCT
TTTGAGAGGAACCAAACAACGTGAGGTGCTGAGGTTCTGGGAATCTCATTGGGTGACACAA
TCCTGGCCATGATTCTACCATTACTCTGCTCTGGGCTCTGATTATGATAGAAGGGAGCCTGG
ACAGACCAAATGATGTCCTGAGAAATGACAACACTCTCAGACCTGTCATGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTGAACACACATCATGGCAACACAGCTTAATACACACT
TTGAGATGGAGGAGTTAAAAAGAAATGTCACAGAAGAACACAAACTTGTTTTATTGGACT
TGTGAATTGGAGTACATACTATGTGTTAGAAATATGAGAAATAAAAATGTTGCCATAAAA
TAACACCTAACGATATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCTGATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTTGAGGCAGCATGGTTGATTAGCATTCCGCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAGGGTGAATTACTCTACCAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTGAATAACTAAACTTTATTAA
CTCAGCGATCTATTCTCTGATGCTAAATAATTATCAGAAACTTCATAATTGGTGA
ACCTAAATGTTGATTTGCTGGTTACTAAATATTCTTACACTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTGTATATAAGCTGTTAAATCTGATAATTCTGATCATT
TTCACTGATAATTGTAAGAATAACCATTGAAAAGGAAATTGTCCTGTATAGCATCATT
ATTTTAGCTTCTGTTAATAAAAGCTTACTATTCTGCTGGCTTATATTACACATATAAC
TGTTATTAAATACCTAACACTAATTGAAAATTACCACTGTCATAGGAATCATT
AGAATGTAGTCTGGCTTTAGGAAGTATTAAAGAAAATTGACACATAACTTAGTTGATTCAA
AAGGACTTGTATGCTTTCTCCAAATGAAGACTCTTTGACACTAAACACTTTTAAAAA
GCTTATCTTGCCTCTCCAAACAAGAAGCAATAGTCCTCAAGTCATATAAAATTCTACAGAAAA
TAGTGTCTTTCTCCAGAAAATGCTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTTGTTTATTCACTGATTAATATACTGTGGAAATTACAGATTATTAATTTTACAA
GAGTATGTTATTGAAATGGAAAAGTCATTTACTGTGTTGTTGATTTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTCATAAAATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTATRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACCTCATCCTCATCCTCTTCTGATAAAGCCCCTACCAGTGCT
GATAAAGCTTTCTCGTGAGGCCTAGAGGCCCTAAAAAAAGTGCTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTCGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCGTGCACCCGTGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCACTGGATGAGTCAGGTCTCCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGCATGGCGGAGATGCCATGCCTACCTCTGCATACAGAAAACCAGTGGAACCC
CGCACCTGTCTGGCTCAATGGCAGGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGAACTGCTGTCTGGAACACCACGGTGGAAAGTCAGGCTTGGGCTGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGCTGCTTCCACGTACTGTGGTCATTATGACATCTGCACGAGGACTGCCATGG
CAGCTGCTCAGATAACCAGCGAGTCACATGCCCTCAGGAACGTGTCAAGGCCAGGACATGCTT
ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAAGTGAAGATCTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTCAAGAAGTGAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTTGCCTTGGATCTGAGAAAGGCTACCAAGTGTGAATGTCCCCGGGCTGGTGT
CTGAGGATAACCACACTTGCCTGGCTGTGCAATCAAATGCCATGAAAGTGAACATCCCAGGG
CTGGTTGGTGGCCTGGAGCTTCCCTGACCAACACCTCCGCCAGGAGTGTCCAACGGCACCCATGTCACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCAAGTGGCAATGACAAGATTGTGGCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGAGCAGCGGGACTCATCATCCGAACCAAGCAAGTGTGATCCCC
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAACTTCGAAACCTCCACTGGAAAT
CATGAGCCGAAATCATGGATCTCCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCAAGCTCGTGAACCCCTACTTGGATTGAGCCGTGGTCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTTGCACCCCACTCCAAAGATCAGCGGGCTGTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTCC
TCCCTGTCTCAAGTTGTGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGTTCTGTGTGGAGTG
TTGGACGGCGTCTCCCGTGTGCCACGGGTTGCCACCGGGAATGCCGTGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCCGATCCGCATCGACTGGAGGACTAGTCGTAGCCATAACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTGGAGCTCTCCCCCACCCTCTAAAGAACATCTGCCAACAGC
TGGGTTCAAGACTTCAACTGTGAGTCAGACTCCACGACCAACTCACTCTGATTCTGGTCATTCAAGTGGCA
CAGGTACAGCACTGCTGAACAATGTGGCCTGGGGTTCATCTTCTAGGGTTGAAAACAAACTGTCCA
CCCAGAAAGACACTACCCCTTCCATTCTCACTTAAACACTCGTGTATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCCAGAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGAAATTATAGTTACCTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAGTGTATCAATAACAGTATATAATT

FIGURE 110

MPPFLLLTCFITGTSVSPVALDPCSYISLNEPWRNTDHQLDSEQPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPEHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYYVRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCASHSCLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LEFLFTNTSCRGVSNGTHVNILFSLKTCGTVVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRSLSYFGIEPVVHV
SGLESLVESCATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQQQLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACCTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGCTCTCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTGGGGAGGACGAGGAGCAGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTCTCGCCTGTTGACAACCTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGATGCGGAACACTCAAGTGGGCCCTGTCCTCAGGCTCCCTGGTCTCCCTGCACTGCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGGTGGGGAGGAGGACATCCTGGACCCCCACTGGTCCCTACGGCA
TCAGCATCCAGTACGACAAACAGCACGTCTGGAGGGAGCATTGGGACTTGGGACTTGGCAG
GCCCACTGCTTCAGGAAACATAACCGATGTGTTCACTGGAAGGTGGGGCAGGCTCAGACAAACTGGCAG
CTTCCCACCCCTGGCTGTGGCCAAGATCATCATCATTAATTCAACCCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCACTCACTTCTCAGGCACAGTCAGGCCATCTGCTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGCTGACATACTGCTGCAGGCGTCAGTCAGGCTATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTCAGGCATCCGGAAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGCATCGTTAGCTGGCTATGG
CTGGGGGGGCCGAGCACCCAGGGATACACCAAGGTCTAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTTAATGCTGCTGCCCTTGCAGTGTGGAGGCCCTTCCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCTCAATTCTGTAAGAGACCCCTCGCAGCCAGAGCGCCAGAGGAAGTCA
GCAGCCCTAGCTGGCCACACTTGGTGTCCAGCATCCCAAGGGAGAGACACAGCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTGAAATGGAAGCAGGCTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTGTGCCAGCCCTGTCCGTCTTCACCCATCCCCAA
GCCTACTAGAGCAAGAACCAAGCTTGTAAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTAAACATCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIP III ALLSLAS II IVVVLIKVILDKYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQMGYSRAVE IGP DQDLDV VEITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPR VVGEEASVDSWPWQVSIQYDKQHVC GGSILD PHWVL TAAHCFRKHTDVF NWKVRAGSDKL
GSF PSLAVAK III IFNPMYPKDNDIALMKLQFPLTFS GTVRPICL PFFDEELTPATPLWIIWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGD SGGPLMYQS
DQWHVVGIVSWG YGCGGPSTPGVYTKV SAYLNWI YNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTAGCAACTAAAAAGCCACAGGAGT
TGAACTGCTAGGATTCTGACTATGCTGTGGCTAGTGCTCTACTCCTACCTACATTAAAATC
TGTGTTTGTCTCTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCAGTGGCTCT
GGCCCAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAACTGTCTCACGTCT
GGAGGCAGTGACTCGGGCAGTGCAAGGTAGCTGAGCCTTGGTAGTGCGGTTCAAGGTGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCGAAGATTCATAGGCATGGCTCCACTGCC
AGGCATCAGCCTGCTGTAGTCATCACTGCCCTGGGCCAGGACGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACCGTCCCCCATCTAACCTTCTATGTCTGCACATCACCTG
ATCCATGGGCTAACCTGAACCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGAAAGGAACCTGTGCCAAATTATGGGTCAAGAAAGATG
GAGGTGTTGGTTATCACAAAGGCATCGAGTCTCTGCATTCACTGGACATGTGGGGAAAGGCTG
CCGATGGCGCATGACACACTCGGACTCACCTCTGGGCCATCAGACAGCGTTCCGCCCGAT
CCACGTACCGACTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGGCCAAAA
TCTGCGATCACCAGCCAGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCAATTCTCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAAGCAAGGCCGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGTGGCTGTAATGCCAACAT
TTGGGAGGCCAGGGCGAGGCTAGATCACCTGAGGTCAAGGAGCTGAGGACTGAGCTGGCCAAACATGG
TGAAACCCCTGTCTCTACTAAAAAATACAAAAAAACTAGCCAGGCAAGGTGTGGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAACATTCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAAATAAAAAAAGAATTA
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLAELTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGCTCTCAGCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAATTGATCCTGTGACCAAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATAACTGG
CATCTACTCGTGGGTCTCAAAAATGTTTATCAAAACCTCAGATTAAGTGAATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAACAGTCAGTG
ATTGGGTCCCAGCAGAAAGCCTATTGAAACCGAGATTCTTCTTAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCCTCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGGAGGAGAAGATCTTCACTTCCGCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAATAGAAATTGATCCCCTGCTGGATGAGAGAG
GTTATTGTTGTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTACTA
GGCTACTACCCATATCCACTGCTACCAAGGAGGAGCTATCTGCTGTCTCATGCCCTG
TAACTGGTGGTGGCCCGCATGCTGGGGAGGGCTTAATAGGAGGTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTCTTATGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKSKICKSLKICGLVFGILALTILFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEIITTFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPLMD
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTAGCTTCACACCTCGCAGCAGGAGGGCGCAGCTCTCGCAGGGCA
GGCGGGCGCCAGGATCATGTCACCACCATGCCAGTGGTGGCGTCCCTGTCATCCTGGGCT
GGCGGGTGCATCGCGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTGGAGGAGCTGCGTGAGGCAGAGTCAGGCTCACCGAATGCAGGCC
TATTCACCATCCTGGACTTCAGCCATGCTGCGAGCAGTGCAGCCCTGATGATGTAGGCATGTCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCATGGTCTCAGGTCTTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACTCTGGATGTCACAGCTAACATGTACACGGCATGGTGG
GATGGTGCAGACTGTCAGACCAGGTACACATTGGTGCAGGCTCTGTTGGCTGGGCTGGTCGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTCATGCCCTGCCGGGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTCAGGCCACAGTGGTGCCTACAAGCCTGGAGGCTCAAGGCCAGCAGTGGCTT
TGGGTCCAACACCAAAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGACTATGTGTAATGCTAAGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCCTAGATTTCTTGTGTTGACTCACAGCTGGAAAGTTAGAAAAGCCT
CGATTTCATCTTGGAGAGGCCAATGGTCTTAGCCTCAGTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTAGCTCACATTTCAATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTTGATGATTTAGACAGACTCCCCCTC
TTCCCTCCTAGTCATAAAACCCATTGATGATCTATTCCAGCTTACCCCCAGAAAACCTTTCTACTGCTG
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGCTCCCCACCCCCACTGGCTAGTAA
ACACTTACTGAAGAAGCAATAAGAGAAGATATTGTAATCTCTCCAGCCATGATCTCGGTTCTT
ACACTGTGATCTAAAGTACCAAAACCAAGTCATTCAGTTGAGGCTCTGAGCTCTCCACTGGAGTCCTCTTCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGAGGTTCTGAGCTCTCCACTGGAGTCCTCTTCTG
CGGGGTCAAAATTGTCCTAGATGAATGAGAAAATTATTTTAAATTAAAGTCCTAAATATAAGTAA
AATAAAATAATGTTAGTAAAGTACACTATCTCTGTGAAATAGCCTACCCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTTGACATTGCTATATGGTACTTTGTAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCTAGCATTGGAGGCTGAGGAGGAAGGATCACTGAGGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCTACAAATACAGAGGAAACATCAGCCAGTC
TGGTGGCATACACCTGTAGTCCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTGAGGCCAGGGAGGT
TGGGGCTGCACTGAGGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTC
AATAAAAATAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGVVAGGLTLIGGVMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAACTCATCCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTTG
GTGGTGGTGGAAATGGGGCACAGTGGCTGTCAGTGTCACTGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTCAAACATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTCCGTGATGTCCTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGGGTGCATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCCTGATTGTTGGAGGAGCTGTTCTGCGTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACTTCCCACGACAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTTAGTTGTATGTTTTAACTTACTATAAGCCATGCAAATG
ACAAAAAATCTATATTACTTTCTCAAATGGACCCCAAAGAAACTTGTATGTTTACTGTTACTGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTGATTCTATAAGCTATTCAGCAGAATGAGATA
TTAAACCAATGTTGATTGTTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATTCT
CTCTTTTATCATTACTTCAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAAAATAGAAACTCAACTATTGCTTTAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTGTTAAAAACAGCTTAGGGATAATGCTTCCA
TTTATAATGAAGATTAAGGCTTAACTCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACCTTATCCTCTCTCCAGAGGCTTTTT
CTTGTGTTAAATTAAACATTAAACGAGATATTGTCAGGGCTTGCATTCAAAGCT
TTCCAGGGCTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGTG
AAAATTTTGTGTTGTATTTGAAGAAGAATGATGCTTGAACAAGAAATCATATGTTGATGGAT
ATATTAAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAAAGAGCAGAAAAATA
TGTCTGGTTTCTTGTACCAAAAAACACAAACAAAAAAAGTTGTCTTGTGAGAAACTTCACCT
GCTCCTATGTGGTACCTGAGTCAAATTGTCATTGTTCTGTGAAAATAATTCTTGT
CCATTCTGTTAGTTACTAAAATCTGAAACTGTATTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCAGCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSIFIENNIIVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSY
RYSIPSHRTTQKSYHTGKKSPSVYRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCGCATTGATGCAGCCTGCGGCGGCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCCTCCGCCTCCAGCTCCGCCTGCCGGCAGCC
GGGAGCCATGCGACCCCAGGGCCCGCCCTCCCCGAGCGGCTCCGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCGTCAAGCGCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCAATGTTACCTGGTACACCTGGATCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGGAAATGTCAGGGAAAGCTTGAGGGAGTCTGGACACCCAACACTAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGGGAGTGTACATT
TACAAAGATGCGTCAAATAGTGTCTAAGAGTTTGTTCAGTGGCTCACTTCCGCTAAATGCA
GAAATGCATGCTGTCAAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTCAGGACCTCTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGGCCCTGAAATGAATTCAACAATTAAATTATCATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGG
TTGGCAGTGTTCAGATTACCCAAAAGGAGATGCTTCACTGGATGGAATTGAGTGTCTCGCATC
ATTATTGAAGAACTACCAAAATAATGCTTAATTTCATTTGCTACCTCTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTAAATAAGTTATGTATAACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAAGTGTGATTTCACACTGTTAAATCTAGCATTATTCAATTGG
CTTCAATCAAAAGTGGTTCAATATTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTGGAATATTGTTGGTCTTTGTTCTCTTAGTATAGCATTGTTA
AAAAAAATATAAAAGCTACCAATCTTGTACAATTGTAAGAATTGTTATATCTGT
TAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAASPQRRLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFGSILRLKCRNACQRWYFTFNGAECGPLPIEAIYLDQGSPEMNSTINIHTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCTGCCTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAAATACATCATGTTTCGATAAGAAGAAATTGAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAAAACTGTAAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTGATGTTGCTGCGAATGCGGTGTTGGGATTATTGTTCTGGAG
TGGTCTCGTGGCTGGCAAGAGATAATGTCCTTGGTCCATCTCCAAGGGTCCAATT
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTCAGCTGACAGGGCTGTCATGCACTG
GCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTGAACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAAGCACTGGTTAGCCCCACTGTCTTACTGACAATG
CTTTCTCTGCCGAACGAGGATGCCCTAAGGGCTGAGGTGTGAAGGCAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTGTCCTTCGCT
ATAACAGCCTCAAAACTTAAGTATAATCAATTAAAGGGCTAACAGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATTGACGAAATGCTTTAATGGAATACGCAACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCTATTCTTAACAATACCTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGCTCTATAATCAGTCGATTCTCTGGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGTCTAACCTCCCTGAGAACCATCCCTGCGAATATTCCA
AGACTGCCAACCTGGAACCTTGGACCTGGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACAATCAATTTCAGCTCAAC
CTGGCCCTTTCCAAGGTTGGTCAGCCTTACTTGCTGGAAATAAAACTAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTACAAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGGTCTTCCAGTGTGTCCTGAATCTGCAGCGCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATCTTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAAGGAAATTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGCTGGCAGCTCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCACTGAAAGAAACTACAGCATCTGCTGGCAAAAGTACTACAGAGAGGTTGATCTGGCCAG
GGCTCTCCAAAGCCGACGTTAAAGCCTAACGCTCCCAGGCCGAAGCATGAGAGCAACCCCCCT
TGCCCCCGACGGTGGGAGCCACAGGCCGCTTCCCTGTCGTCATCCCTGCTGGTTAT
CTACGTGTCATGGAAAGCGGTACCTGCGACGATGAAGCAGCTGCAAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGACATCCTAAGCAATGACTCCCAGCACCAGAATTATGTA
GATTATAACCCACCAACACGGAGACCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGTGGTCTTTATTGAACTCTGGTGAATCAAGGGAACGCGATGCCCTCCCTCCCC
TTCCCTCTCCCTCACTTGGCAAGATCCTTCCTGTCGTTAGTGCATTATAACT
GGTCATTTCCCTCTCATACATAATCAACCCATTGAAATTAAATACCAATCAATGTGAAGCTT
GAACCTCCGGTTAATATAACCTATTGTATAAGACCTTACTGATTCCATTATGTCGATTT
GTTTAAGATAAAACTCTTCTAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNVAGMIRLKELHLEHNQFSKLNLAFLPRLVSLQNLQLQWNKISVIGQTMSTWSSLQRDL
SGNEIEAFSGPSVFOCVPNLQRLNLDNSNKLFIGQEILDWSISLNDISLAGNIWECNRNCSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVLVLLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCCCGAGGAGGAGAGGCTTGTCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAAACTCGACTCCAATGCAAGGAGAAGCAGCTTGTCTC
GGTGGGAGACGGTCAAGAGAATCTGCCCCATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGAACACCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCAATGGTACATATGAACATCTCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCCTTGGAAATCAGTCATTGGAGGGATGATGGTGGTATTGGCCAGTTT
TTAGCCAATCCAAC TGACCTAGTGAAGGTTAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGGCAGGCTGGGTACCAATATAAAAGAGCAGCAGTGGTAATATGGGAGATTAA
ACCACTTATGATACTGAAACACTACTTGGTATTGAATAACACCAGTGGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTCACTGAGTCTATATAAAGGCTTTTAC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSrimnQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCAGCGAGGCGGACGGCGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCAGCGGCTGGCGCTGGTACGGGGCCTCGGGGGCATCGGCCGGCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTCAGGCTACCCGGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTCAGGCAGCACCAGTGGTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCCTGACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTAA
CCCCTGTCGTGACCCACTCTATAGTGCACCAAGTATGCCGTACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCACGTCTCAGCACCCCCCGACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTGCCCTGCCTCTGGATTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTCCTGTCCACACCCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGCCCTGGCACTTGCCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCCAAATCCCCATCTTGTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTTATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCCTGCCCACTGCACCCCTCTCCCCCTATCTATCTCCTCTCGGCTCCCC
AGCCCAGTCTGGCTTCTGTCCCCCTCTGGGGTCACTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRILALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININMSGHRLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGCCTTCCTCAGCCTCCTGCCGGTGCCTACACCAT
CATGTCCTCCCACCCCTCCTTGACTGCGGGCGTCAGGTGCAGAGTCTCAGTTGCCGGAGC
ACCTCCCTCCCAGGGCAGTCGCTCAGAGGGCCTGCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGTAAGGCCATGGAACCTTGGTGAATCACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGGAATGTTAGGTCGTTCTGTCGCTGTTCAATTCAAGTAGCCACAGCCACCTGTGG
CCGTTGAGTGCTGAAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCTCATTTATTAA
TTAATTCTTAACTGATAGTTGACATATTGGGGTACATGTGATATTGGATACTGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCTACTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCACCTCCAAGTAGCTGGACTACAGGCAT
GCACCACAATGCCCAACTAATTGTATTTAGTAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTGAACTCCTGGCCTCAAACAAATCCACTTGCCCTGCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCTGGCTAAACATTATCTTTCTTGTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGACTTCTCCCTCT
ATCTAACTGTATATTGTACCAGTAACCACCGTACTTCATCCCCACTCCTCTATCCTCCCC
AACCTCTGATCACCTCATTCTACTCTCACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCCTTCTGCTGGCTTATTCACCTAACATAATGACTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMPPFKRVFCQDGTVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCCTCATGGGACTAAAATGGCTGTTCCCTCAAGTAGCACCTATCAGTTATGGCTAAATCCCTG
TCCATCTGTGTGTCGCTGCGATGGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACTCTACCTTCAGAACAAACAAATAATGCTGGATTCCCTCAGAT
TTGAAAAACTTGTGAAAGTAGAAAGAATATAACCTATACCAACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTCCCTGTCCTGAATCACCTAGCACAATTCCCTGGGTTT
GCCCAAGGACTATAGAAGAAACTACCGCTGGATGATAATCGCATATCCACTATTCATCACCACATCTTCAG
GTCTCACTAGTCTAAACGCCTGGTCTAGATGAAACCTGTTGAACAATCATGGTTAGGTGACAAGATT
TTCTTCAACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGAATTCCCTGACTGCTGCACCAGTAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATGGGTGCCCAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGTCATAAACTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAATGATTCTCGCAACAATCCCTGGTATTGGTGCAGATGAAATG
GGTACGTGACTGGTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTATGTGCCAAGCCCCAGAAA
AGGTTGGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGATTGTAAGC
ACCATTCAAGATAACCACGTGCAATACCAACACAGTGTATCCTGCCAAGGACAGTGGCAGCTCAGTGAC
CAAACAGCCAGATAATTAAGAACCCAGCTCAACTAGGATCAACAAACACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTGTGATACCATTCATATCTCTGGAAACTGCTCACCTATG
ACTGCTTGAGACTCAGCTGGTTAAACTGGGCCATAGCCCGCATTTGATCTATAACAGAACAAATTG
AACAGGGGAAACGCACTGGTACAGTGGTACAGGCTCACAGGCTGATTCCCTATAAAAGTATGCATGGT
CCATGGAAACCAAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCT
CGAATGTACAACCCCTACAACCAACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAACCCCAATTAC
TTGGCTGCCATCTGGTGGGGCTGTGGCCCTGGTTACCATGGCTTCTGCTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAACGCTGGCACTAAGAAGGACAACCTATCCTGAAATCAGGAAACTCTTTCAAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGGAGTTGTAATACACACCATACTTCCCTCTAATGGAATGAATCTGT
ACAAAACAAATCACAGTGAAGGAGCTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVRCDAAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKEHLHQENNIRTITYDSL_SKIPYL
EELHLDNSVSAVSTEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNVLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQOTTGSPSRKTITITVKSVDTSIHWKLAALPMTALRLSLWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCVPMETSNLYLFDETPVCETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTI FPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCCCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACCTGTCTTATGTTGCCAGGGGAGGTGGGAAGGAGGTGGAGGAGGGCG
TGCAGAGGCAGTCTGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGCTGCTCCTGCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGCTGGACACTGCTGAAAGTAAAGCCACATTGCAAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCGAGTGTGGAAGAGCAGCTAAAAGTGTCCGGGAGAAGTGGGCCAGGAGCCCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCAAAGTACCTAACAGAGAGTTCCAGCTGACCCCTCCAGCCCG
GTTTGAAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTGGTGTACCCACGTTGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCTGTCACCAACTGCTCTTCTCTGGGCCAGAATGAGGGATGCACACAGG
GACCACCTCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGAGCTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCCTACCCGGACATCTTCATGGAAAACATCATGTTCTG
TGGAAATGGGCGGCTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCAGAAAAACAATTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCCTGGTTCAAGC
AATTCTTGCCTCATCTCCCGAGTAGCTGGACTACAGGAGCGTGCACCACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACCTCTGAT
CTCAAGAGATCGGCCACCTCAGGCTCCAAAGTGTGGGATTATGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCTAGATGGCTGCTCCACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACTCCTGGCAGAACACCCCCCAGCAAACAGAGAGGCCACCCATCCACACCG
CCACCAAGCAGCGCTGAGACGGACGGTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTAGTCTCATCCCTAGATCCTGGAGGGACGGATCACATCTGGAGAAGAGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAACTTGGAGGCCAGGGTAGGTGGAGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGGRRACRGSILGLARAQGAERVTSSEQRPMASLGLLLLLTLAPPLWSSSLPGLD
TAESKATIADLILSALERATVFLQRLPEINLDGMVGRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSILHYLKLSDPKYLREFQTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDI FMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGCTCTTCCGCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTGGAGGTGCCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTCAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGCGGCCGAGTGCTGGTAGACGGAGAAGACGACGTGGTTCTTAAG
ACAGATGGGAGTTTGTGGTTCATGATATACCTCTGGATCTTATGTAGTGGAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTGAGTGGATATCACTTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTCAGGT
CCACCTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTCCTTATTGATATTGTGCTTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAAACAGGCAAAAGTGGGCTGGCAAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAACTT
GAGTCATCCCGACGTTGATCTTACAATGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATCTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTATAGAAAATGCCATTAATAAAATTAT
ATGAACTACTATACTATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 136

MAAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPVMMMLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCTCCCCAGTCCCCGTGGGTCTGAGGGGA
CCAGAAGGGT GAGCTACGTTGGCTTCTGAAAGGGGAGGCATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTTCACGGGAGGCTGGCAGT
TTTCTTACTCCTGTGGTCTCCAGATTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTGCTGCGTTTATCTCCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTTCTGAGATAACGGGCAGTG
TGCAAGCAAAGATGGAAACATTGACATCAGAAATCTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAAGATCAGCAGCCTCGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTGTGAAGGC
TTTGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAAGAGGAGGCATGCCAA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGTCACAGTGTATCTTATTATGCATTACTTG
CTTCCTTGATGATTGTCTTATGCATCCCCAATCTTAATTGAGACCATCTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTTATTGCTATT
ATGTATTTATTTTACTGGACATGAAACTTTAAAAAAATTACAGATTTACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAATTGACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGATGTGAAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATATTGTGTATCTTCCAG
CCAGGAATCCTACACGGCCAGGATGTATTCTACAAATAAGTTCTTGCATACCAAAAAAAA
AAAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLD RVFKNYQTPDHYTLRKIS
SLANSFLT IKKDLRLSHAHMTCHC GEEAMKKYSQLSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAAGCGGGCTGCAGCAGGGCAGGGCTCCAGGTGGGGTGGTCCGCATCCAGCC
TAGCGTGTCCACGATCGGGCTGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAGGGTC
TAGGGATCGCGGTCTTCTTCTGGGGATTCTTCCCGGCTCCCGTGGTCTGCCAGAGGGAA
CACGGAGCGGAGCCCCAGCGGGAAACCCCTGGCTGGAGCCAGTCTAAGTGGACCACGCTGCC
ACCACCTCTTCAGTAAAGTGTATTGTTCTGATAGATGCCCTGAGAGATGATTTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTTGAGGGAAAAAGGAGCAGTCAAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGAG
CCTTCTGGCTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTCGTCAGATTACAGAGGT
GGATAATAATGTACGAGGCATTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAACTC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTCAGGGCCAAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCACTGCAAGGAGAGA
GACGCCCTTACCCAAATTGCTGGTTCTTGTTGACCATGGCATGTCATGAAACAGGAAGTCACG
GGGCCTCTCCACCGAGGGAGGTGAATACACCTCTGATTTAATCAGTTCTGCGTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCTAGACGGATGTCGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCATTCAGTTGTTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATGAAAAAGATCCTGGTTGAGCAGTTAAATGTCAGAAAGATTGCAATTGG
GAACGTGATCAGACTGTACTTGGAGGAAAGCATTAGCAGAAGTCTTCAACCTGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAAGACGCTGAGCTTGTCCCTGAGTGACAAAGTGGCC
TTCTCACCTGCTCTGCTCAGCGTCCCACAGGCACTGCAAGAAAGGCTGAGCTGGAAAGTCCA
CTGTCATCTCTGGTTTCTGCTCTTATTGGTATCTGCTGGTCTTCCGGCGTTCACTG
CATTGTTGTCACCTCAGCTGAAAGTTCGTCACCTCTGTCGCTCTGTCGCTGGCGCAGGCT
GCCCTTCTGTTACCGACTCTGGTTGAACACACTGGTGTGCTGCAAGTGTGCCCCGGAC
AGGGGGCCTCAGGGAAAGGACGTGGAGCAGCCTTATCCCAAGGCTCTGGGTGTCCCACACAGGTG
TTCACATCTGTCGCTGTCAGGTCAAGTGCCTCAGTTCTGGAAAGCTAGGTTCTGCACTGTTAC
CAAGGTGATTGAAAGAGCTGGCGTACAGAGGAACAAGGGCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCACAGGAGCTGGAGCAGCTGCAAGTGTGCCCCACTCATCTGCCACCCCC
CACTCAGGGAGGGTCAAAAGGAGACTGGTGTGCAAGGACTCATCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCAAGGGACGTTCTGTTGGAATTCTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGGAGTGGTGGTGAAGGCACTGAAGAAGAGCGGATGGTCAC
ACTCAGATCCACAGAGCCAGGATCAAGGGACCCACTGCACTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCCCTGCACAGCCCTCATCCCTCTGGCTTGAAGCCCTGAGGGCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCACTCAGGGCACAGGCTCTCGGAGGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLPPPLF
SKVVIVLIDALRDDFVFGSKGVFKMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDLILYLHHLGHDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTCCAGTTATCGTACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT
TCCCTATAGAAAACACTGCCAGCACCTAAGACCACTCAACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACTACATACGCCAGAGATCTTCTTGCAATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAGTTTGTCCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTATAGGGCTCAGGTGGCTCCTGGA
ACATGCTGGAGTCGGCGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTCACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAACTGCCATTGAACGCCCTCGCTA
ATTGAACTAATTGATAAAAACACCAACCTGCTCACT

FIGURE 142

MLLLLLEYNFPPIENNQHLKTTHTFRVKNLNPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGCTTCCCTGCTCTGTAGGAGTGCCTGCTGCCAGTGTGGGGTGGAGACAAGTTG
TCCCACAGGGCTGCTGAGCAGATAAGATTAAGGGCTGGGCTGTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGCCTGCCGTGG
AAGGGAGGTCTGCTCTGGCGCTGCTGCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTGAATTCCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTGAAGAGCAGACATTGACAACACTGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTCACCTGCTTCTTCACCATCAGCACCAAGGCCCTGGATGACTCAGTTCA
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTAGTTAGAAAATCC
ACACATCTGAGCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGGCACTGTCATGCTGCTTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCCAGGAGCCTCATCAAGCAAGGCCAGCAGCTGACATT
GACTTCGGAGGCCAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGGCGCGAACCCAGGGGGAGTCCAGAACGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTGGAGAGGGCGCAGGACT
TCGGGTACCATGCACCAAGCCAGTGCCTCTGCCTTGGCTTGTGATCTGGCTATGGTGAAAAT
AAGCTTGGCAGGAGGCTGGCAGTACAGAGGCCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCCTGGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCCACGTATGCGCCCTGGTATGCGCTGCCTGCTGATAAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCAGTGCAGGCCACCTGCTTCACAGTACTTCCAACAACCTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTTGTCTAACCCCAAGGTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGTACTTATGTCTCACAGGGACCCACATCCAACATGTAT
CTCTAAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWILATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACATGTCTGGTCCAGAGTCTCATT
CCTGATGATTTAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTGGCCTCCTCT
CTGTCTTCTTCCCTTTCTTCTTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGTAGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
CAACTTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCC CGCGCCCCAACCCCTGCTTATCCCTGACCGTCGAGTGTCAAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCCGCCCCACACCCACCCCTCGCTCTTCTGTTTACTCCTCTTTCTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAGAAGAATGGGTT
CCTCGGGACCGGCACCTGGATTCTGGTGTAGTGTCTCCGATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAGTCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACATATCCTCCAGAAAACAAGCCAGGTCAAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAACTCTCAAGGAAGCCACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTGCTAAGGGAGAAAAGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTTCCAAATTCTATGCGCTACTGAAAAGTATTGATTCAAGAAAAGAGCAAAAGAGAAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTCAAATATGGAACAAATATCT
CCAGAAGAAGGTGTTCTACCTTGAAACTGGATGAAATGATTGCTCTTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCAGCACCCTAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTGAAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCTTATGGAAAGC
CATCAGAAAAAATATTGAATGGTGAAGAAAACATGACAAAAAGGAAATAAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGATTATAGCAGCCTGAAAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAATTGGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTTAAACATAGCTTCTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTPPENKPG
QSNYSFVDNLNLLKAITKEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVTPEKTPMAAIQDGLAKGENDETVSNTLTLNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLKNATDNISKLFPAPEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCTGAGTGGGGCGCTGTGCTTCCATGAGGACTCGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTCAATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGGCTGTCCCCCGTCATCCTGGGTGTCAGGGTGAAG
CCAGTGCCTGTCAATGTGGGTGGGGCAGGGAGCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCAGGAAATCCAAGAGCTTACCCCTCACCGGCGGACATGGGCTCACC
TCCAGCTCGAGTCGGCTGCCTACCCGGCTGTTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAACTCACCCAGCTTCCCAGAATGGTGGCTGAAATGCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGAGAGGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCCCTCCCCACTGGATGGTGCTACTGCTGTGAAACCTTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCGCAGGCCAACCCATCCCCAGTTGAGCCTATA
GGGTCACTAGCTCCACATGAAGTCTGCACTCACCCTGTGCACTGAGGAGAGGGAGGTGGTATA
GAGTCAGGGATCTATGGCCCTGGCCAGCCCCACCCCTTCCCTTAATCCTGCCACTGTCTATA
TGCTACCTTCTATCTTCCCTCATCATCTTGTGTTGGCATGAGGAGGTGGTGTGAGTGTAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTTAATCCATGAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAAGTGGCATATAATTGCAATTAAATTAAAGATAACCTATTATATATT
TCTTTATAGAAAAAAAGTCTGGAAGAGTTACTTCATTGTAATTGCTGAAATGTCAGGGTGGCAAGTAT
AGGTGATTTTCTTTAATTCTGTTAATTATCTGTATTCTCTTAATTCTACATGAAGATGA
ATTCCCTGTATAAAAATAAGAAAAGAAATTATCTGAGGTAAAGCAGAGCAGACATCATCTGAA
TTGTCCTCAGCCTCCTGAGTAAATTGCAATTGAAATGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGTAGTCAAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTGT
GTGGCTGGAATCTGGGTAAGGAACTTAAAGAACAAAAATCATCTGGAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGATTGGATCCAGTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCTCTGAAATTCACTCCTCTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCCTGCAAGATGTAGTTAGTAAAGACAAGGTGATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTTGTATGAAAAGAGGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAGCCACAGCTAAGAAACACCAAGGATTGTGCAACC
ATCAGAAGCTTGGAGAGGCAAGAAGAATTCTCCCTAGAGGTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAAACGAGAAAGAATAAATTCCGGCTGTTTAA
GCCACCAAGGATAATTGGTACAGCAGCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTGTTCACATTCTGTGTGTCCTCCCTCCACAAATGTACCAAAGTTGTCTTGTGACCAA
TAGAATATGGCAGAAGTGTGGCATGCCACTTCAAGATTAGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGCCCCACCCCACTCATCTTGGCTCACTCGCTCTGGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGGCTATAAGAGACTACGGTAAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATGATTTGTTGTTTAA
AAGTTGCTCAGTTGGCTAACTGTTATGCAAGCAATAGATAAATATGCAAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEELSVVNRWLDASLSPVILGVQGGS
QCLSCGVGQEP TLLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCTGGCCACCAGCTGCCTCCTCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCGCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTC
CAGCCTTATATGCAGGAGGTGGTGCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGAAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCCTAGTTACAAAGGAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATAATTGATAACATTCAATTGTAACGGTGTTC
TATACACAGAAAACAATTATTAAATTGTCTTTCCATAAAAAGATTACTTTCCAT
TCCTTAGGGGAAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATTATAAGACTGCATTTATTATCATTATTATAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVLIGEKLFHGVSMERCYLMKQVLNFTLEEVLFQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CACTCAGTCCCCGACTGTGACTGAGGTGCGAGTGCAGTACAGGTCAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCAAGC
TGCCAGGTTGGGCTGGGGCAAGTGGAGTGAGAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTCACAGGTGGTGCAT
TCTTGCAATGGTCATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGACTCCTGTAGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTGAAACCGCTCCCCCAGGACCTGTACCAACGCCGT
TGCCTGTGCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAAACAGACTGTCTCTACAGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGCTGTGGAGCAG
CAGGATCCCGGGACAGGATGGGGCTTGGGAAACCTGCACTCTGCACATTGAAAGAG
CAGCTGCTGCTAGGGCCGCCAGCTGGTGTCCGTCACTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTTCCCTTCCATCCCTGCTACCTG
GCCCACAGGCACATTCTAGATATTCCCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTTGTTACTCATCACTCAGTGAGCATCTACTTGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGAAGCTGTTATTGAATGTATAGAGATTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATGTCGCTCGT**GCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTGCAACAGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGCAGATGCCAGCATCCGTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGAAAAGCAACTTCCAGTCACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTGGTGGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTCATTGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTACCTCACCAAGGCTGCCAGACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAACGCTGTGGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAAGTGAACCTCACAAACCACTCCCCTGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATCGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGAATGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTGCCACAAACAGGCGTCCCTTCCCT
GGATAACAAACAAAAGCAAGCCGGAGGCTGGCTGCCTCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGGAGGCACGAAAGGATCAAGAAGACTTCTT
TCTACCACCAACTACTGCCCTTCAATTAGGTTCTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACCAATTGAGAAGTGGCTCATCCTGAAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTCAGTGGCTGCCACTAAAAGAAGGCA
GCAGACAAAGTCGCTTCCCTTCAATTGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCTTAAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGGCTACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTAGTGTCTGCCCAAGTACCAACCTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTSVATGDYSILMNWSWV
LRADASIRLLKATKICVTGKSNSFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGSPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNNK
SKPGGWLPLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristylation site.

amino acids 116-122

Amidation site.

amino acids 488-492

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCC~~T~~GCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCGCAGCTCGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTGGGACCCCAACGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCC~~T~~GGTCGTCCGGAGGAAGCACCAAGGCTGCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLSILGLAFLSEAAARKIPKGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVMSMRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVFQLEKVLVTVGCTCVTPVIIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACAGGGCTGACTGGGGTGTCTGCCCCCTGGGGGGGGCAGCAC
AGGGCCCTCAGGCTGGGTGCCACCTGGCACCTAGAAGTGCCTGTGCCCTGGTTCTTGCTGTCC
TGGCACTGGGCCGAAGCCCAGTGGTCTTCTCTGGAGAGGCTGTGGGCCCTAGGACGCTACC
CACTGCTCTCCGGCCTCTCTGCCGCTCTGGACAGTGAACATACTCTGCCCTGCCCTGGGACAT
CGTGCCTGCTCCGGGCCCTGCTGGCGCTACGCACCTGCAGACAGAGCTGGTGCTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTGCTGCGTGTGGCTGTCCACTGGCGTGCATGGGACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCAAGTCGTGCTCTCCAGGCCTACCCACTGCCCCCTGCCGTCTGC
TGGAGGTGCAAGTCGTGCTGCCCTGTGCAAGTTGGTCAGTGTGGCTGTGGTATATGAC
TGCTTCAGGGCTGCCCTAGGGAGTGAGGTACGAATCTGGCTCTATACTCAGCCAGGTACGAGAA
GGAACATCAACCACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTCAGCAGATGGTGACA
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTCGGCCCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCCTAAACCCGGTGGCACAAAAACCTGACTGGACCGCAGATCATTACCTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTCAAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCCGCAGACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCTGCAGAGCTGGCTGCTGGACGCACCGTGCCTGCCGCAGAACGGCACT
GTGCTGGCGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCCCTGGGAGAACG
TCACTGTGGACAAGGTTCTGAGTCCCATTGCTGAAAGGCCACCTAACCTCTGTGTTCAAGGTG
AACAGCTGGAGAAGCTGCAGCTGCAAGGAGTGTGCTTGTGGCTGACTCCCTGGGCCCTCTAAAGA
CGATGTGCTACTGTGGAGACAGCAGGGCCCCCAGGACAACAGATCCCTCTGTGCCCTGGAACCCA
GTGGCTGTACTTCACTACCCAGAAAGCCTCCACGAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCCCTCGTGTGGCTGCCCTACTCTTGCCG
CTGCGCTTCCCTCATCCTCTCTCAAAAGGATCACCGAAGGGTGGCTGAGGCTCTTGAAA
CAGGACGTCCGCTCGGGGGCGCCCGCAGGGGGCGCAGGCTCTGCTCCCTACTCAGCCGATGA
CTCGGGTTTCGAGGCCCTGGGGGCCCTGGCGCTGGCCCTGTGCCAGCTGCCGCTGCCGTGG
CCGTAGACCTGTGGAGCCGTGGTAAGTGAAGCGCGCAGGGGCCGTGGCTTGGTTCACGCCAG
CGGCCAGACCCCTGCAGGAGGGCGCGTGGTGGCTTGCTTCTCTCCCGGTGCCGTGGCCT
GTGCAGCGAGTGGTACAGGATGGGTGTCCGGCCCCGGCGCACGGCCCGCACGACGCCCTCC
GCCCTCGCTCAGCTCGTGTGCCCAGCTTCTGCAAGGGCCGGCCCGCAGCTACGTGGGG
GCCTGCTTCGACAGGCTGCTCCACCCGGACGCCGTACCCGCCCTTTCCGCACCGTGCCGTCTT
CACACTGCCCTCCAACTGCCAGACTTCTGGGGCCCTGCAAGCAGCCTGCCGTGGCTCC
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCAGGCCCTCAGCAGCCCTGGATAGCTACTC
CATCCCCCGGGGACTCCCGCGCCGGACCGGGGGTGGGACCAAGGGCGGGACCTGGGGGG
CGGGACTTAAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSCRILWSDSDILCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVLFSQAYPTARCVLLEVQVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSAADGDNVHLVLNVEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLTLQSWLADAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKLEFPPLLKGHPNLCVQVNSSEKQLQLQECLWADSLGPLKDDVLLLETRGPQDNRSILCALEPSGCTSILPSKASTRAARLGEYLLQDLQSQCQLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSADDGFERLVLGALASALCQLPLRVAVDLSRRELSAQGPVAWFHAQRRQTLQEGGVVLLFSPGAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVFTLPSQLPDFLGLQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCCCTGAGGACCCCTGGATCTGCTCCAGCAGTGAATTCCAGTCAGCAACTTGA
AAACATCTGACGTGGGACAGGGCCAGAGGGCACCCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGAAAGAAGGGTGTAGCGGATCACCGGAAGTCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGTACCGCT
GTCAGTGCAGGGAGGCCGGTCAAGCACCAGATGACTGACAGGTTCACTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATGATTGAGATGATTGTTCATCTA
CCCCCACGCCAACCTCGTGCAGGCATGGCCACCGCTAACCTGGAAAGACATCTTCAATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTGGAGGGAAAGCAGAGA
ATATGAGTTCTTCGGGCTGACCCCTGACACAGAGTTCTTGGCACCACATGATTGCGTCCCCA
CCTGGGCAAGGAGATGCCCTACATGTGCCAGTGAAAGACACTGCCAGACCGGACATGGACC
TAATCCTTCTCCGGAGCTTCTGTTCTCATGGGCTCTCGTCCAGTACTCTGCTACCTGAG
CTACAGATATGTCAACAGGCCCTGACCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCGCTGCGCTTCATCACAGAGCTCCAGTGTGACCCCTGTTGACCTCAGCGGCCCCAGC
AGTCTGGCCAGGCTGTCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCTAGCTGTCGAGATCACCTACTTAGGGCAGCCAGACATCTCATCCTCCAGC
CCTCCAACGTGCCACCTCCAGATCCTCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCCATCTATGCACTCAGGTACCCCCGAAGCTCAATTCCATTCTACGCCAAC
GGCCATCTAAGGTCAGCCTCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGTATGCAAGGTTCTGGAAAGACTCCCCACTGGGACACTTTCTAGTCTAA
CACCTTAAGGCTAAAGGTCAAGCTTCAAGAAAGGCCACAGCTGGAGCTGCAATTGAGGCT
TTCTGCAAGGAGGTGACCTCTTGGCTATGGAGGAATCCAAGAAGCAAATATTGCAACAGC
CCCTGGGATTGCAACAGACAGAACATCTGACCCAAATGTGTCACACTGAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCTCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTGCAACCTCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGG
TGCTGGAGTCCCTGTGTGTCAGGATGAGGCAAGGCCAGGGCTGAGACCTCAGACCTG
GAGCAGCCCCACAGACTGGATTCTTCAAGGCTGGCCAGCTGAGTGGGAGTCC
AGGGGAATGGGAAAGGCTGGGCTCTCCCTACCCAGTGTCACTACATCTGGCTGTCA
ATCCCATGCTGCCATGCCACACTCTGGCATCTGGCTCAGACGGGTGCCCTGAGAGAAC
AGAGGGAGTGGCATGAGGGCTGGCATCTGGGAGGAGACAGGCAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTGGGAGCAGCTGTGAGACAAGCGCTGCTGGCTGAGCCCTG
CAAGGCAGAAATGACAGTGCAGGGAAACTCCGAGGTCCAAGGCCACCTC
CTAACACCATGGATCAAAGTCTCAGGGAAATTGCCCTCCCTGCCCTGTCAATTGTT
ACAATCTAGCTGCACAGAGCATGAGGCCCTGCCCTGTCAATTGTT
ACAAGGAGGCTGGCAGAACAGAACAGAACAGAACAGAACAC
ACTTCTGCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGTGTGGCCTGAGCTCA
TTCCCAAGCCAGGGCAACTGCCCTGAGCTGACGATTTCACTGCTTCTGAGAACAAAGC
GAAATGCAAGGTCCACCAGGGAGGGAGACACACAAGCCTTCTGCAAGGCCAGGAGTT
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAACAC
TGTACTGATGTCAACACTTGCAGCTCTGCCCTGGGTCAAGCCATCTGGCTCAAATTCCAGC
CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTT
ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGT
TCTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCTAAACGGTAGCTATTTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGQRITRKSCNLTVETGNLTLYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNQVRVLTFOQPLRFIQEHVLIPIVFDLSPGSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILOPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGTGGCCACAACATGG
CTGCGGCGCCGGGCTGCTCTGGCTGTCGTGCTGGGGCGCTCTGGTGGTCCCGGCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTTGAAGACTTCACGGGCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGCTACTACAAACTGGCAGGGGATCCCTGAACTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCCAAAAGATTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCGCTGAGCCCGAGGCATTCAAGAGCTGATTCAGGAGGATGGAGAAGGTGCTTCT

CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAAGCGAAGCAGAAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGAGA
AGACAGATGCTTACAAAGTCCTGAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCGTATT
CATTACAGCAAAGGATTTCGTTGGCATCAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTTGAGCCAAACAAAAATATATTATTTCCTTCTAAGTA
AAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVCGLDEECMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAAEVSQHREKSPPEESRGRELDPVPEPEAFRADSEDEGA
FSESTEGLQGQPSAQESHPHTSGPAANAQGVQSSLDTFEEILHDKLKPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACCGCGACCCCTCGAAGTCTGAACCTCCAGCCCCGACATCCACCGCGGGCACAGG
CGCGGCAGGCAGGTCCCGGCCGAAGGCATGCGCAGGGGTCGGGCAGCTGGCTCGGG
GGCGGGAGTAGGGCCGGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGGGCTGCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGCGCCGCTGCTCTGGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GCCAAAAGGTGTGTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTCAGGAGGCACGCCCTGGCTTGAGAGTGAGGAGGAGTCCCTCA
GCCTTGAGAATGAAGCAGAACAGAACAGATTAATAGAGAGCATGTTGAAAACCTGACAAAACCCGG
ACAGGGATTCTGATGGTGTAGTTCTGGATAGGGCTTGGAGGAATGGAGATGGCAACATCTGG
TGCCTGCCAGATCTTACCAAGTGGTGTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCTGCCGAAGTGTGAAAGTGTGATGTATACCAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTACCAAGTGGATGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTTGCTTTGGAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAATAGTCCAACCCAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAACTATAAAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTGGATGGCTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTGTTATATGTCTATTATTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTGTATCAACACGTCGGGAGTA
TGTGTGTAGAAGCAATTCTTTATTCTTACCTTCATAAGTGTATCTAGTCATGTAA
TGTATATTGTATTGAAATTACAGTGTGCAAAGTATTACCTTGCTAAGTGTGATAAAA
ATGAACTGTTCTAATATTATTGTCATCTCATTTCATAACATGCTTGTGATAAG
AAACTTATTACTGTTGTCACTGAATTACACACACACAAATATAGTACCATAGAAAAGTTGT
TTTCTGAAATAATTCACTTTCAGCTCTGCTTTGGTCAATGTCTAGGAAATCTTCA
AATAAGAAGCTATTCAATTGTCAGAACATGTGCTTAAACACTCAAACATTACTAGAGGCAAGGAT
TGTCTAATTCAATTGTCAGAACATGTGCTTAAACACTCAAACATTACTAGAGGCAAGGAT
TTGTAATAATGTAACTTTGTTAATAGGTGCATAAACACTAATGCAATTGAAACAAAAGAAG
TGACATACACAATATAATCATATGTCCTCACACGTTGCTTAAACATTGAAAGCAGCTCTGA
GGGTTCTGAAATCAATGTGGCCCTCTTGCCCCACTAAACAAAGATGGTTGTCGGGTTGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATAAAAGAGGTATGTGGTTGAGACCAGGGTAATAGTCACATCAGTGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAACTACGAAATCGTGTGAAAATGGTTGG
AACCCATCAGTGCATATTCAATTGATGAGGGTTGCTGAGATAGAAAATGGTGGCTCTT
CTGTCCTATCTCTAGTTCTCAATGCTTACGCCCTGTTCAAGAGAAAGTTGTAAC
CTGGTCTCATATGTCCTGCTCCCTTAACCAAATAAGAGTTCTGTTCTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSSLGAALLCGHGAFCRRVSGQKVCADFKHPCYKMAFHESRVSFQEALACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGFHWIGLWRNGDGQTSACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217